Olfactory short term memory: Understanding perceptual representations of odours and the role of encoding strategies in working memory.

Andrew Gordon Moss

A thesis submitted in partial fulfilment of the requirements of Bournemouth University for the degree of Doctor of Philosophy.

July 2017

Copyright Statement

Abstract

This project examined the representation of odours in working memory. There is a paucity of research examining specific olfactory working memory ability, and there are equivocal findings concerning the availability of an internal representation to consciousness and the extent and influence of verbal coding. This thesis first describes the creation of a comprehensive database of odour normative data, which contributes to the future control and manipulation of olfactory stimuli in experimental research. Individual differences were assessed across these odour ratings, and four dimensions identified as suitable for future stimulus control. Olfactory working memory was then examined using the n-back task with verbalisable and hard-to-verbalise odours. A working memory advantage for verbalisable odours was replicated (Jönsson et al., 2011), but this advantage was unrelated to the adoption of a verbal rehearsal strategy. Instead, effects from a concurrent rotation task provided tentative evidence for an attentional refreshing process. Controlled working memory processes were shown to be reduced for low verbalisability odours, though there was no evidence in a remember-know task for a switch to more automatic processes (i.e. familiarity). However, in an individual-differences analysis of multi-modal n-back performance, only low verbalisable odours were unrelated to verbal and visual working memory. The n-back working memory findings may therefore reflect a perceptual memory that is unavailable to consciousness, and an important role of semantic information in the generation of an internal representation that can be manipulated in working memory. Finally, this thesis provided a first examination of item-specific proactive interference effects in a memory task, which showed absent proactive interference effects for low verbalisability odours and which supports mediation of an olfactory representation through odour verbalisability. It was suggested that a 'fuzzy' representation for low verbalisability odours results in a weak link between an item and a conflicting familiarity signal.

Status of manuscripts from this thesis

The work in Chapter 2 has been published as: Moss, A. G., Miles, C., Elsley, J. V., & Johnson, A. J. (2016). Odorant Normative Data for Use in Olfactory Memory Experiments: Dimension Selection and Analysis of Individual Differences. *Frontiers in Psychology*, *7*, 1267.

The work in Chapter 5 has been published as: Moss, A. G., Miles, C., Elsley, J. V., & Johnson, A. J. (2017). Item-specific proactive interference in olfactory working memory. *Memory*, 1–15.

List of contents

List of figures

List of tables

Acknowledgement

There are many people to whom I am indebted in completing this thesis. To my supervisor, Dr. Andrew Johnson, thank you not only for providing steadfast guidance throughout my research, but also for your enthusiasm and encouragement; whether it is persuading me to present at conferences, explaining that those reviewer suggestions aren't so difficult after all, or telling me to take the weekend off to play frisbee.

To my second and third supervisors, Dr Jane Elsley and Dr. Chris Miles, thank you for your interest, critique, and discussions, they are greatly appreciated. I would also like to thank Martin Tomkins for his invaluable technical knowledge.

Becca, Rhiannon, Mary, Chris, Darren, and the others in P104, thank you for providing an environment where we can share our successes, worries, anecdotes, cake, and beer.

To Becky, thank you for your unwavering support and patience, and to the rest of your family for their kindness and reinforcement.

Finally, thank you to my own family, for your endless positivity, excitement, and care.

Chapter 1: Introduction

1 Chapter Summary

The ability to remember olfactory information has been largely ignored in general models of memory, with a focus instead on the verbal and visual domain. However, remembering odours is essential for functions that include food detection and identification, hazard avoidance, and social communication (Stevenson, 2010). This memory is typically discussed in terms of the Proustian phenomenon, where an odour will evoke a distinctive and detailed memory (Chu & Downes, 2000; Herz & Cupchik, 1995). However, there are also requirements for 'pure' odour memory, where a memory for an odour is required without association to an experience. For example, a smell might be discriminated in a same/different task (Jehl, Royet, & Holley, 1995; Rabin, 1988), or recognised as having been experienced previously (e.g. Jehl, Royet, & Holley, 1997; Zucco, 2003). Of particular interest in this thesis is the ability to temporarily retain and manipulate olfactory information within working memory, where there is a paucity of work. Several every-day tasks, including comparing odour intensities, the freshness of foods, or simply distinguishing between two different smells, all presumably require some temporary olfactory representation that must be maintained in a memory store (White, 2012). Furthermore, attention to the components of odours is necessary to establish whether the mixture is tainted by some contaminant (Dacremont & Valentin, 2004; Thomas-Danguin et al., 2014).

This introduction considers whether unique features of olfactory memory can be accommodated within established models of working memory. It begins by discussing evidence for the representation of odours as a perceptually-based memory trace, and the influence of other-modality representational codes. It is important to establish the extent to which an olfactory memory task actually reflects olfactory memory or, instead, memory for a re-coded version of the stimulus (e.g. verbal labels). This is followed by an introduction to an object-centred account of odour perception under the *olfactory-centred unitary model* (Wilson & Stevenson, 2006), which ties olfactory perception to activated odour objects in long-term memory. The historical and current conceptualisations of odour memory are then discussed in terms of the unique features olfactory memory displays compared to other modalities. That is, whether memory for odours should be considered qualitatively distinct to memory for other modalities is discussed, with reference to forgetting functions, serial position effects, and the ability to consciously access stored olfactory representations (e.g. Stevenson, 2009; Zucco, 2003). Research is then discussed that examines olfactory working memory, with consideration of modular and unitary models of working memory. Proposals for an independent 'olfactory buffer', analogous to verbal and visuo-spatial working memory slave systems, are also discussed (e.g. Andrade & Donaldson, 2007).

1.1 Odour Representations

1.1.1 Perceptual and verbal representations of olfactory stimuli

Prior to discussing models of memory, it is important first to consider how odours may be represented within memory. Similar to other non-verbal stimuli (Melcher & Schooler, 1996; Pickering, 2001; Schooler & Engstler-Schooler, 1990), the extent of identification, naming, and verbal re-coding is clearly an important consideration in discussions of olfactory memory ability, though findings are equivocal over whether verbal processes form an essential part of olfactory memory (e.g. Jönsson, Møller, & Olsson, 2011; Murphy, Cain, Gilmore, & Skinner, 1991; White, Hornung, Kurtz, Treisman, & Sheehe, 1998; Wilson & Stevenson, 2006). Whilst removing all influence of semantic information and/or verbal labelling from explicit olfactory memory is likely impossible (White, Møller, Köster, Eichenbaum, & Linster, 2015), the use of dual tasks (Andrade & Donaldson, 2007), and manipulation of familiarity and verbalisability (e.g. Jönsson et al., 2011; Møller, Wulff, & Köster, 2004), have been used to elucidate the processes involved and to assess whether a perceptually-dominated memory for odours is possible.

Above-chance recognition ability has been shown for unnameable, unidentified, or unfamiliar odours (e.g. Cleary, Konkel, Nomi, & McCabe, 2010; Jönsson et al., 2011; Kärnekull, Jönsson, Willander, Sikström, & Larsson, 2015; Miles & Hodder, 2005; Møller et al., 2004; M. J. Olsson, Lundgren, Soares, & Johansson, 2009; Yeshurun, Dudai, & Sobel, 2008; Zelano, Montag, Khan, & Sobel, 2009). Memory performance for unidentified odours (i.e. a label given that was not a match to the veridical name, nor a close miss) was well above chance in Olsson et al. (2009), and odours difficult to verbalise have shown above chance performance in the n-back task (Jönsson et al., 2011). In Møller et al., (2004), explicit recognition ability was shown for unfamiliar odours. In this study, superior performance for younger (compared to older) adults was attributed to differences in working memory ability rather than an increase in the use of verbal labelling (indeed, one might expect vocabulary to increase with age, and therefore if labelling improved memory one might predict superior odour memory in older groups). Furthermore, in a free recall task (in which participants were required to recall the names of odours presented at encoding), response confusions for odours are far greater for perceptually similar odours (e.g. lemon and orange), than for perceptually dissimilar odours where their verbal labels are phonologically similar (e.g. lemon and melon) (White et al., 1998). Together, these findings are evidence for a perceptual code in the encoding of an odour, which can support a perceptually-based odour memory that is not reliant upon semantic or verbal mediation.

It has been suggested that recognition memory may be possible when an odour is identified in the perceptual sense (e.g. matched to a stored odour engram, see below), but unidentified semantically (Stevenson & Mahmut, 2013b). For example, *recognition without identification* (RWI) describes the phenomenon of discriminating target items from lures despite being unable to identify the test item (e.g. Cleary, 2002; Peynircioğlu, 1990). This is demonstrated by having participants study a list of stimuli, and at test presenting the stimuli with features occluded (e.g. as word fragments with letters removed, Peynircioğlu, 1990). They are then required to identify the word and rate whether it was presented in the prior list. It is these ratings that support recognition ability without identification. For words, this is explained by the use of perceptual information such as abstract orthographic information, or phonological information, which gives rise to the feeling of familiarity to enable recognition (Cleary, 2002). The presence of RWI has also been shown for picture fragments (Langley, Cleary, Kostic, & Woods, 2008), auditory word fragments (Cleary, Winfield, & Kostic, 2007), and odours (Cleary et al., 2010). In Cleary et al. (2010), participants were presented with odours and names at study, and at test were required to name the target and lure odours and rate the likelihood that each item was presented previously. Like other modalities, recognition was shown in the absence of identification. Importantly however, when only the odour name was presented at study the recognition ratings from presentation of an odour did not discriminate targets and lures without identification. This suggests that odour RWI is a perceptually-driven phenomenon, because it is necessary that the perceptual experience of an odour is experienced at encoding. It should also be noted that demonstrations of olfactory RWI in a typical recognition paradigm may be under-reported, due to the reluctance of participants to respond 'old' to an item when unable to report the identity of the odour (R. A. Frank, Rybalsky, Brearton, & Mannea, 2011).

Other evidence for olfactory perceptual memory has come from studies of implicit memory. Degel, Piper, and Köster (2001) examined implicit memory for odours by incidentally presenting odours within different rooms and then later asking participants to provide a judgment of 'fit' for various odours when shown pictures of different rooms. They showed implicit episodic memory for odours and the contexts (rooms) in which they were perceived (by previous odour-room pairings receiving a higher 'fit' rating), and even demonstrated proactive and retroactive interference effects when multiple odour-room pairings were experienced (Köster, Degel, & Piper, 2002). Importantly, Köster et al. reported that the effects were removed when the odours were identified, suggesting that these memories were formed implicitly without conscious identification/labelling of the odour. Based upon this observation, Köster et al. argue that such implicit memory tasks are a more ecologically valid assessment of odour memory, with identification or recollection of odours less important than change-detection and hazard avoidance in olfactory perception (Köster, 2005). The implicit memory work of Köster and colleagues therefore demonstrates that olfactory memories can be formed without verbal identification/elaboration.

In summary, the literature supports an olfactory memory that is not uniquely attributable to the recoding of information as a verbal or semantic code, or which requires identification for memory to be successful. Instead, the results support a memory ability that is, to some extent, reliant upon the perceptual representation of an odour (e.g. White et al., 1998). However, olfaction is not uniquely underpinned by perceptual code; there is a strong influence of familiarity, semantic information, and labelling, which must be considered in any model of olfactory memory.

1.1.2 Additive or facilitative effects of verbal labelling

Wilson and Stevenson (2006) categorise verbal influences on olfactory processing by their effects on both perceptual and non-perceptual processing. That is, being presented with a choice of names for an odour, or maintaining a name between study and test, may result in activation of the stored representation assigned to that odorant to be enhanced (see below for an object-centred account of olfactory processing; Cessna & Frank, 2013; De Wijk & Cain, 1994b; R. A. Frank et al., 2011; Wilson & Stevenson, 2006). This results in perceptual benefits (e.g. discriminability, Rabin, 1988) because the increased activation allows the direct activation of an appropriate odour object. Alternatively, non-perceptual benefits may occur by providing additional codes through which an odour can be represented in memory (Andrade & Donaldson, 2007; Annett & Leslie, 1996; Paivio, 1990; Stevenson & Mahmut, 2013a). That is, additional codes may switch the way an odour is represented into something more easily encoded and retrieved (Herz & Engen, 1996), or result in a dual memory trace (Paivio, 1990).

To examine the role of labelling in olfactory memory, some studies have sought to encourage labelling and to examine subsequent memory performance. Such an approach has produced mixed findings. For example, some olfactory memory research has shown no facilitative effect from effortful verbal labelling (Engen & Ross, 1973; Lawless & Cain, 1975; Zucco, 2003), nor effects from verbal interference tasks undertaken during the retention interval (Engen, Kuisma, & Eimas, 1973; Zucco, 2003). This apparent independence from verbal elaboration is described in Zucco (2003) as evidence against a consciously accessible representation of odours, because such a representation is impervious to interference from other information, and also to facilitation from additional encoding (Zucco, 2003). However, other studies have shown impaired olfactory memory when performing concurrent verbal tasks (Annett & Leslie, 1996;

Murphy et al., 1991), and a large number of studies have shown improved olfactory memory under circumstances where verbal elaboration is possible (R. A. Frank et al., 2011; Jehl et al., 1995, 1997; Jönsson et al., 2011; Kärnekull et al., 2015; Lyman & McDaniel, 1986, 1990). This advantage is particularly apparent for familiar odours that are easily or consistently named (R. A. Frank et al., 2011; Kärnekull et al., 2015; M. J. Olsson et al., 2009), and indeed identifying an odour will produce similar recognition levels and recollective experiences to verbal stimuli (M. J. Olsson et al., 2009). Odour familiarity is not however necessary for facilitative effects of verbal labelling to be observed, as a recognition advantage has also been shown for unfamiliar odours when subjects are given or generate their own verbal label at encoding, an effect observed across both short (20 minutes) and long (24 hours) delays (Jehl et al., 1997). These separate but likely related effects of familiarity and ease of naming, coupled with evidence of a detrimental effect from a concurrent verbal task during encoding, suggests an influence of semantic information in odour memory.

The above evidence indicates that mapping an odour percept to semantic information is advantageous for recognition memory. Similar to memory for other stimulus types, a general elaborative network model (Anderson, 1983), or dual-coding model (Paivio, 1990), may explain an advantage from verbal labelling. Both models suggest that multiple encoding processes will aid recognition; elaboration by providing multiple retrieval paths that increase the likelihood olfactory information is retrieved (Anderson, 1983; Bradshaw & Anderson, 1982), and dual-coding through functionally distinct traces in verbal and non-verbal systems which contain referential links that allow one code to activate another (Paivio, 1990). In dual-code theory, these verbal and non-verbal codes are proposed to be statistically independent, and additive in their facilitative effects on memory (Paivio, 1991). That is, each code may differ in terms of memory ability, but the inclusion of the other will improve performance.

A dual-coding proposal is similar to the proposed separation of information from verbal and visuo-spatial systems in working memory, and their interaction as a bound representation in the episodic buffer (Baddeley, 1996, 2000; see Section 1.4.1). However, odour recognition memory improvements occur with both verbal and pictorial elaboration (Lyman & McDaniel, 1986, 1990), and can be impaired by concurrent verbal and visual tasks (cf. Andrade & Donaldson, 2007; Annett, Cook, & Leslie, 1995; Annett & Leslie, 1996). Whilst it should be noted that the findings in Lyman and McDaniel (1986, 1990) have been criticised due to inferences made from the reduction of false alarms rather than from changes in explicit recognition (Zucco, 2003), the results support a dual-code interpretation beyond simply employing secondary verbal code (Herz, 2000). Specifically, an advantage from visual imagery supports the presence of multiple traces, rather than being confined to a simple distinction between verbal and non-verbal memory (Annett & Leslie, 1996; Lyman & McDaniel, 1986). In summary, a memory advantage from learned associations to odours appears to be due to creation of a multi-modal dual-trace, and not the specific inclusion of verbal labelling per se.

An alternative consideration to dual-code theory has proposed that verbal coding will supersede an olfactory representation when semantic information becomes available (for example, when remembering familiar odours, or when provided with a name) (Herz & Eich, 1995; Herz & Engen, 1996). This is comparable to other modalities, for example, where recognition memory for word fragments that are identified benefit from encoding tasks that focus on meaning (Cleary, 2002). When fragments are not identified, no benefit occurs and recognition is suggested to rely on the use of abstract visual information. A similar process may occur for odours, where identification changes the way items are represented in memory (Cleary et al., 2010; Zelano et al., 2009). Some supporting evidence for an encoding switch for odours comes from priming effects, where an odour and its name presented as a prime will produce the same speed improvement in a subsequent odour-name matching task as a name-only prime (Schab & Crowder, 1995). This suggests that facilitative priming effects from odour presentation are due to the availability of a verbal code. In short, proponents of an encoding switch suggest odours are recoded into verbal representations, and thus the memory may cease to be a representation of the odour percept but instead reflect verbal memory for that odour label (White et al., 2015).

However, several findings are difficult to accommodate within an encoding switch explanation. Andrade and Donaldson (2007) showed memory for common odorants (which presumably were identifiable to some extent, though this was not reported) was not impaired by a concurrent verbal task, and suggests verbal processes are not an intrinsic part of olfactory memory. Furthermore, as mentioned previously, there is strong evidence for perceptual confusions when odours are perceptually (compared to when phonologically) similar (White et al., 1998). Importantly, this task required naming responses, suggesting a strong perceptual representation in olfactory memory even when verbal labels are required at test. The findings of White et al. therefore suggest that the perceptual representation has primacy over the verbal label.

It is therefore prosaic to argue that olfactory information can be stored in memory with a dual code (Lyman & McDaniel, 1986; Yeshurun et al., 2008; Zelano et al., 2009), but this secondary code may not be specific to verbal labelling (Annett & Leslie, 1996). This is consistent with a proposal in Baddeley (2012) that odours may be processed within the episodic buffer, suggesting a working memory that acts upon bound representations of olfactory and other modality information. However, behavioural and neuroimaging data remains equivocal on whether this dual representation occurs in all instances of olfactory short-term memory, or whether odours can be represented by solely perceptual representations (See Section 1.4.1 for discussion of an independent olfactory buffer; Andrade & Donaldson, 2007; Yeshurun et al., 2008; Zelano et al., 2009). A model of olfactory perception and memory has been presented however, discussed in Section 1.1.3, that attempts to consolidate these findings (Wilson & Stevenson, 2006).

1.1.3 An object-recognition approach to olfactory perception

The olfactory-centred unitary model (Stevenson & Boakes, 2003; Wilson & Stevenson, 2006) describes olfactory perception in terms of object recognition, where odours can be perceived from within a complex olfactory environment as a unitary whole (Stevenson & Wilson, 2007). The model suggests perception of a distinct odour object is the result of activation of a stored olfactory representation in long-term memory, and is therefore reliant on prior perceptual experience with the odour. This activated representation is also linked in a two-way relationship with semantic information, where top-down knowledge can facilitate activation of the stored olfactory representation, and semantic information can be activated following presentation of an odour and co-opted for use in memory (Stevenson & Boakes, 2003).

This section first describes the proposed physiology of olfactory perception that leads to the selection of a distinct odour object. Perception of an odorant begins when volatile chemicals bind to a distinct combination of receptors in the olfactory epithelium (Buck, 2004), and this pattern of activation projects to the glomerular layer of the olfactory bulb before being conveyed to regions that constitute the primary olfactory cortex (including the piriform cortex, Gottfried, 2010). This activation pattern is based on the spatial orientation of receptors activated, and temporal information related to the speed with which chemical components interact with the receptors (Buck, 2000). However, odours that are both the target of perception, and present in the background, will bind to these receptors equally. Thus a unique activation pattern upon presentation of every smell will be generated, and will make identification impossible. Consequently, there is a requirement for the odour to be separated from background odours (i.e. figure-ground separation) to allow perception of a discrete odour object (Gottfried, 2010; Stevenson & Boakes, 2003; Stevenson & Wilson, 2007; Wilson & Stevenson, 2006).

The olfactory-centred unitary model is an integrated model of odour perception and memory because mnemonic processes are proposed to drive the ability to perceive this distinct odour object (Stevenson & Boakes, 2003; White, 2009; Wilson & Stevenson, 2006). Two key processes are described: habituation and pattern matching. That is, habituation of background odours at the level of the odour receptors and olfactory bulb neurons allows the piriform cortex to analyse only pattern features related to the odour object (M. E. Frank, Goyert, & Hettinger, 2010; Stevenson & Wilson, 2007; Wilson & Stevenson, 2003a). This is demonstrated in piriform activations from a target odour following adaptation to stable background odours that matches activations when the target is presented without the background odours (Kadohisa & Wilson, 2006).

The second proposed process describes the matching of this input pattern to stored perceptual object templates (*engrams*, Stevenson & Boakes, 2003). Pattern-matching is possible because an odour that has been experienced previously will have an odour template contained within an *object store* (Stevenson & Wilson, 2007; Wilson & Stevenson, 2006). When the odour is experienced again there is a close match between its input pattern and this stored template, which will be activated strongly amongst a small number of other similar olfactory templates (Stevenson & Mahmut, 2013a). The strong activation of a limited number of similar templates, combined with inhibition of other templates due to this activation strength, results in perception of a discrete odour that is separable from the background and other odorants, and gives rise to the perception of the odour's quality (Stevenson & Boakes, 2003; Stevenson & Mahmut, 2013a; Wilson & Stevenson, 2006). In contrast, a novel or unfamiliar odour will not have an exact match in this store, and consequently will activate many stored templates to a much lesser extent. Novel and unfamiliar odours therefore produce a vague perceptual experience, and an unstable percept once the stimulus is removed (Stevenson & Mahmut, 2013a). This novel odour is also less discriminable from other unfamiliar odours because it is redolent of many odour representations (Mingo & Stevenson, 2007).

Association of semantic information to an odour template can be effortful, and strengthened through repetition (De Wijk & Cain, 1994a; Rabin, 1988), though an association between odour and other information at encoding may also occur incidentally (e.g. Degel et al., 2001). When experiencing an odour, the activation strength of the odour template will increase the likelihood that this associated link to episodic or semantic knowledge is excited (Stevenson & Boakes, 2003), resulting in identification of the stimulus, generation of a name, and other contextual information. It should be noted however that olfactory identification is typically very poor, and participants overestimate their ability to name odours (Jönsson & Olsson, 2003; Jönsson, Tchekhova, Lönner, & Olsson, 2005). The reason for this is debated (Stevenson & Mahmut, 2013b). Naming may be difficult because of the 'fuzzy' activation that occurs as the result of multiple template activations within the object store (Jönsson & Olsson, 2012), meaning that unless an odour is identified (the correct odour template is activated), the correct name cannot be retrieved. This is supported in demonstrations of 'tip-of-the-tongue' states, which shows a strong feeling-of-knowing in the absence of a name, but unlike other modalities is accompanied by very little semantic information (Jönsson & Olsson, 2003; Jönsson et al., 2005). That is, the semantic information is not available because the odour object has not been identified. In comparison, Stevenson and Mahmut (2013b) suggest an odour may be identified perceptually, but naming is poor due to a weak link between semantic and perceptual systems. This is perhaps because there is rarely a requirement for an interaction between the two systems (because an odorous object can usually be seen, and hedonic information may be processed separately, Stevenson & Mahmut, 2013b).

Naming ability can be drastically improved by allowing the selection of a name from a multiple choice list (e.g. Cain, 1979; De Wijk & Cain, 1994b; R. A. Frank et al., 2011). Providing a choice of names serves to cue the correct response, and is suggested to lower the threshold for activating the appropriate odour template (R. A. Frank et al., 2011). This not only facilitates identification (Stevenson & Boakes, 2003; Wilson & Stevenson, 2006), but also make odours more discriminable (Rabin, 1988), and appear more intense (Distel & Hudson, 2001). This account can also explain the facilitative role of verbal labelling in olfactory memory (e.g. Lyman & McDaniel, 1986). For example, it is possible for verbal codes to independently support olfactory recognition by providing an additional code by which olfactory information can be maintained (Herz, 2000; Stevenson & Mahmut, 2013a). The facilitation can therefore occur because retrieval of a verbal code effectively mimics the effect of cued identification of the odour, lowering the threshold of activation for that odour engram when it is re-presented. Alternatively, odours that are consistently identified will consistently activate associated semantic information that includes its verbal label, and later presentation of the same olfactory stimulus will activate the same object, and with it, the remembered verbal label (R. A. Frank et al., 2011). This could mean that only the verbal label needs to be maintained over the retention interval, though such an explanation is inconsistent with the evidence against an encoding switch, discussed in Section 1.1.1 (e.g. Andrade & Donaldson, 2007). In summary, this model suggests therefore that how an odour is represented may vary depending on the ability of an odour to be identified perceptually, and the subsequent ability to activate semantic information related to that odour.

The above model has subsequently been elaborated to explain odour memory. Wilson and Stevenson (2006) suggest that odour memory is a direct consequence of olfactory perception. That is, recognition memory is proposed to occur as a consequence of residual activation of the stored odour templates, which facilitates the reactivation of the representation when the odour re-occurs (Stevenson & Boakes, 2003; Stevenson & Mahmut, 2013a; Wilson & Stevenson, 2006). This is proposed as a unitary system, so does not differentiate between short and long term memory (but see White, 2009, and Section 1.2.1 for a discussion). Furthermore, whilst this activation allows performance in short-term recognition memory tasks, proponents of this model suggest the activation of an odour engram does not reflect a conscious representation (Stevenson, 2009; Stevenson & Mahmut, 2013a). This may reflect similar processes to those described in RWI (Cleary, 2010), where odours can be recognised in some familiarity-based system without identification.

However, the above model, describing memory driven by unconscious access to stored perceptual representations, has been criticised by its inability to explain findings where directed attention is used to aid odour detection and identification (White, 2009, 2012). A general attentional model of olfactory processing has been proposed (e.g. see White, 2012, and Section 1.2.4 below), though a requirement of this model is the ability to process odours in working memory. The allocation of attention for olfactory imagery or rehearsal in working memory may not be possible, however, if there is an inability to access a conscious representation of perceptual olfactory representations (see Section 1.2.3 and 1.2.4). The next section considers these issues, and other demonstrations of olfactory memory that are not analogous to other-modality systems, with the eventual purpose of considering how odours may be accommodated within a working memory system.

1.2 Unique features of olfactory memory

Section 1.1.1 presented putative evidence for memory of an olfactory perceptual code, and evidence that whilst this representation can be facilitated by verbal processes it is not intrinsically dependent on them. Employment of a perceptual representation in olfactory memory task (rather than relying upon verbal recoding) indicates the existence of an odour specific buffer. Further support for the existence of such an olfactory store would be evidenced by qualitative differences in olfactory memory relative to other stimulus modalities. To this end, this section considers how olfactory memory is comparable to memory for other modalities. That is, despite some similarities to visual and verbal memory, there are clearly unique characteristics of olfactory memory that support the existence of a separable system, different to verbal or visual memory (see Herz & Engen, 1996).

1.2.1 Forgetting functions: the existence of a dissociation between STM and LTM

A dramatic early claim was that olfactory memory was not separated into STM and LTM components (Engen et al., 1973; Lawless & Engen, 1977). A key finding underpinning this claim is poor initial recognition ability for odours, coupled with a flat forgetting curve across both short intervals up to 30 seconds (Engen et al., 1973; F. N. Jones, Roberts, & Holman, 1978) and long intervals between one day and twelve months (Engen & Ross, 1973; Lawless & Cain, 1975). This finding contradicts the canonical logarithmic forgetting curve reported for verbal (Ebbinghaus, 1885; Murre, Dros, Gais, Born, & Dey, 2015) and visual (Roger N. Shepard, 1967) stimuli. The apparent absence of temporal effects on the forgetting of olfactory memories contradicts the traditional distinction of a highly accurate yet fragile STM and a less accurate but more stable LTM (e.g. Atkinson & Shiffrin, 1968) and suggests that olfactory memory may be qualitatively different to other modalities.

However, rather than interpret the flat forgetting function as evidence for an absence of olfactory STM, an alternative explanation for this apparent lack of 'decay' concerns how odours are represented in memory. Odour representations were characterised as a distinct unitary event with little attribute redundancy (Engen, 1982; Herz & Engen, 1996). That is, for odours, unlike other stimuli, there are a lack of other hierarchical features which form that item (Collins & Loftus, 1975). Consider, for example, visual objects which may be characterised by multiple features that include shape, colour, size, and location (Allen, Hitch, Mate, & Baddeley, 2012); odours, in contrast, comprise a single feature. This lack of additional features that comprise the odour representation is proposed to make the initial level of recognition poor (Engen et al., 1973), but also results in minimal retroactive interference through a reduced number of features that can overlap across odours (Lawless & Engen, 1977; Schab, 1991; Zucco, 2003).

Despite these initial differences, more recent studies have reported similar forgetting curves for named odours, unidentified odours, and verbal stimuli (M. J. Olsson et al., 2009), albeit from different initial recognition levels (see also, Kärnekull et al., 2015, for similar forgetting functions to that of faces, and Köster et al., 2002, for similar cross-modal forgetting effects with implicit odour memory). Moreover, there exists indirect, yet compelling, evidence for odours being susceptible to retroactive interference shown via the negative effects of a same modality suffix (Miles & Hodder, 2005; Miles & Jenkins, 2000) and same modality secondary task (Andrade & Donaldson, 2007; Walk

& Johns, 1984). Consequently, it appears odour memory may be susceptible to interference from other olfactory stimuli in a way that is analogous to the retroactive interference observed in verbal and visual memory (see Chapter 5 for more detailed examination of interference in olfactory memory).

The existence of olfactory STM is proposed due to the sequential nature of most olfactory memory tasks, such as the ability to distinguish the smell of one odour from another. These discriminability tasks presumably require the temporary maintenance of the removed stimulus until it is ready to be compared to the newly presented item (White, 2009, 2012). However, though influential modular accounts of memory suggest a dissociation between STM and LTM (e.g. Baddeley & Hitch, 1974), temporary storage may instead be explained as a consequence of residual activation of representations in long-term memory (Cowan, 1999; Wilson & Stevenson 2006). White (1998, 2009) considered in two reviews whether a STM-LTM distinction is appropriate for olfactory memory, with specific consideration of the olfactory-centred unitary model. White (1998, 2009) examined three areas of evidence for STM-LTM dissociation; (1) capacity differences, (2) coding differences, and (3) neuropsychological evidence. These areas of evidence for STM-LTM dissociation are considered in detail below.

In respect to capacity differences for STM and LTM, short-term memory for odours is affected by the set size (i.e. the number of to-be-remembered items) (Engen et al., 1973; F. N. Jones et al., 1978), and the effect of set size is comparable to that reported for verbal stimuli (Murdock, 1961). Furthermore, concurrent task findings support a limited-capacity buffer consistent with a modular working memory system (Andrade & Donaldson, 2007). That is, Andrade and Donaldson showed a disruptive effect of a secondary olfactory task on odour memory performance, suggesting a limited capacity for processing olfactory information in a short-term memory task. In comparison, Zucco (2003) showed no detrimental effect of a task performed in a longer retention period (10 minutes), requiring participants to rate olfactory, acoustic, or visual stimuli. This suggests a functional difference between olfactory STM and LTM. However, Stevenson (2009) argues that perceptual similarity between the task items can explain the above effects of set size and concurrent olfactory processing. In the object-centred unitary model of olfaction, this would result in greater overlapping activations of stored representations, and detrimental performance from additional odours in the task as a result (Wilson & Stevenson, 2006). An additional source of evidence for olfactory short-term memory is the strong recency effects observed in serial memory (A. J. Johnson & Miles, 2007; Miles & Hodder, 2005; Reed, 2000; White & Treisman, 1997), though it should be noted that findings of extended recency (e.g. A. J. Johnson & Miles, 2007; cf. White & Treisman, 1997) might instead be interpreted within the olfactory centred unitary model as a reduction in the likelihood that olfactory features would overlap (see Section 1.2.2 for further discussion of differences across serial position functions).

In respect to coding differences for STM and LTM, it is suggested that short-term coding is modality-specific whereas long-term memory is supported by a more general semantic code (Baddeley, 1996). Like capacity differences, support for this distinction can be found with modality-specific interference in a short-term memory task (Andrade & Donaldson, 2007). Specifically, Andrade and Donaldson (2007) found that performance on a yes/no recognition olfactory task was disrupted to a greater extent by a secondary yes/no olfactory recognition task undertaken during the retention interval, compared to the effects of a secondary verbal or visual task. Furthermore, in an examination of olfactory LTM, Zucco (2003) observed no disruption to odour memory from an olfactory interference task, which may suggest long-term memory for odours is not represented in a modality-specific code (though they instead interpret this finding as evidence against a conscious representation in memory, see Section 1.2.3). It should also be noted that not all studies have shown coding differences. For example, both perceptual and phonological coding has been observed for olfactory short-term memory (White et al., 1998), and both perceptual and semantic coding has been observed in olfactory long-term memory (Cain, De Wijk, Lulejian, Schiet, & See, 1998).

In respect to neuropsychological evidence, classical evidence for a STM-LTM dissociation has been shown with hippocampal lesion amnesics who exhibited intact memory at retention intervals of 5-minutes but deficits at lengthened intervals of 1-hour (Levy, Hopkins, & Squire, 2004; cf Dade, Zatorre, & Jones-Gotman, 2002, who reported an impairment to olfactory recognition that was unaffected by retention interval). Furthermore, short term odour recognition has been shown to be impaired in epileptic patients (Carroll, Richardson, & Thompson, 1993), though in this study the extent to which long-term olfactory memory remained intact was not examined. Evidence from H.M. (an amnesic patient with a deficit in consolidation to long-term memory, Eichenbaum, Morton, Potter, & Corkin, 1983) is often cited as one half of a double-dissociation for a STM-LTM distinction (Milner, Corkin, & Teuber, 1968), and in a battery of tests of olfactory functioning H.M. showed a deficit in odour quality discrimination (Eichenbaum et al. 1983). This suggests odour discrimination is dependent on intact representations in long-term memory, and may be evidence against a functionally separate olfactory STM (although alternatively, it is possible that the olfactory pathways for odour discrimination were damaged in addition to long-term memory, White, 1998).

In summary, White (2009) argued in favour of the olfactory-centred unitary model, suggesting that whilst short-term memory for odours is similar to other modalities, it (and other modalities) may not necessarily exist as architecturally separate systems (Cowan, 1999; Stevenson & Mahmut, 2013a; White, 2009, 2012; Wilson & Stevenson, 2006). Indeed, more generally, the dissociation of STM and LTM remains controversial in the memory literature (e.g. Bhatarah, Ward, & Tan, 2006; Bjork & Whitten, 1974; Tan & Ward, 2000). However, the evidence described here is far from conclusive against such dissociation, and other modalities have shown more convincing evidence for a distinction between the two types of memory. These apparent differences between olfaction and other modalities provide evidence against a general unitary process across modalities (see a detailed discussion of unitary and modular systems in Section 1.4).

1.2.2 Serial position effects

Serial position effects refer to the differences in memory performance as a function of the position of items in the original sequence. That is, whether relative to other items in a list, memory is superior at certain positions in the lists (first, last etc.). The extent to which olfactory memory differs to other modalities can be assessed by comparing the serial position functions cross-modally (although in cross-modal comparisons one should be cognizant that methodological differences can alter the shape of the serial position function, Ward, Avons, & Melling, 2005). The patterns of recognition memory for odours, assessed through 2-alternative forced choice (2AFC) recognition, typically produce functions consistent with other stimulus types (A. J. Johnson & Miles, 2007; Miles & Hodder, 2005; cf. Reed, 2000). However, whilst primacy effects may indicate the rehearsal of information early in a list (Rundus, 1971; Tan & Ward, 2000, although see Tan & Ward 2007, for critique of a rehearsal account of primacy), these have been absent for olfactory versions of order memory tasks that are typically characterised by primacy (A. J. Johnson, Cauchi, & Miles, 2013; A. J. Johnson & Miles, 2009). Indeed, when primacy has been observed for odour sequence memory, this has been attributed to verbal labelling of the odours (Annett & Lorimer, 1995; Miles & Jenkins, 2000; Reed, 2000). However, it is worth emphasising that Ward et al. (2005) show that the traditionally reported cross-modal differences (between visual and verbal memory) can be explained by task rather than stimulus differences. At present, there exist insufficient studies that compare the serial position functions for olfactory and visual/verbal memory when both stimulus types are applied to the same task demands. Initial studies (A. J. Johnson & Miles, 2007, 2009) have produced conflicting findings suggesting that olfactory memory may be similar to other stimulus types in respect to item but not order memory. However, more comprehensive work in this area is needed.

1.2.3 Imagining an odour

Mental imagery is proposed to directly engage perceptual pathways to recreate the experience of an item, in the absence of a presented stimulus (Farah, 1988; Kosslyn & Thompson, 2003; cf. Pylyshyn, 2003). Imagery is closely related to working memory (Tong, 2013), and consequently to directed attention, consciousness, and the ability to actively rehearse or refresh olfactory information in memory (M. R. Johnson, Mitchell, Raye, D'Esposito, & Johnson, 2007; Stevenson, 2009). The re-creation of an olfactory perceptual experience would therefore support an ability to access an internal representation that is analogous to processes in visual memory (Kosslyn, Ganis, & Thompson, 2001), but findings in this area are equivocal (e.g. Crowder & Schab, 1995; Rinck, Rouby, & Bensafi, 2009; Royet, Delon-Martin, & Plailly, 2013; Stevenson, 2009; Tomiczek & Stevenson, 2009). However, although investigations of olfactory working memory are limited (see White, 2012, for a discussion of olfactory working memory, and Section 1.3 in this thesis), evidence for such an ability (e.g. Dade, Zatorre, Evans, & Jones-Gotman, 2001; Jönsson et al., 2011; Zelano et al., 2009) suggests some capability to both image and consciously access an internal representation of odours.

This section explores the nature of olfactory imagery, and examines how such an ability might differ to that seen for other modalities.

It is clear that the ability to imagine an odour is poor in comparison to visual, auditory, or even haptic stimuli (Herz, 1996, cited in Herz, 2000; Stevenson & Case, 2005). For example, odour images are self-reported as less vivid and more difficult to produce than other modalities (Ashton & White, 1980). Indeed, some research suggests such an ability may not be possible at all (Crowder & Schab, 1995; Herz, 2000). For example, a mental image is proposed to re-create the perceptual experience of an odour, so paired-associative memory would be expected to show similar performance regardless of whether cues at learning and test were the actual stimulus or a prompt to imagine the stimulus. Herz (2000) showed this was not the case, as cued recall was impaired when the cue switched from an *imagine* format at study to the actual odour at test, compared to when both study and test were *imagine* cues. The presence of switch effects for olfactory memory was interpreted as an inability of imagery to reproduce the perceptual experience of an odour. However, it should be noted that there was no alternative modality presented in this task with which to compare these effects of cue-switching.

In contrast, several avenues of research have supported an olfactory imagery ability (for reviews see Rinck et al., 2009; Stevenson, 2009; Stevenson & Case, 2005). Olfactory hallucinations, for example, have been reported in the absence of a sensory stimulus, and is evidence for an ability to recreate the perceptual experience of an odour (Stevenson & Case, 2005). Furthermore, like other modalities, participants imagining the experience of an odour will activate overlapping brain regions with those when actually experiencing an odour (Bensafi, Sobel, & Khan, 2007; Djordjevic, Zatorre, Petrides, Boyle, & Jones-Gotman, 2005; see Kosslyn et al., 2001 for a discussion of the neural basis of imagery for other modalities). Similarly, activation related to odour processing in the inferior frontal gyrus occurs in anticipation of olfactory task demands, and continues in the short-term memory period beyond removal of the stimulus (Rolls, Grabenhorst, & Margot, 2008).

Tomiczek and Stevenson (2009) demonstrated generation of an odour image through facilitative effects of odour imagery priming in an odour-name association task, though these imagery effects were dependent on the ability of participants to name odours in an earlier task. In three experiments, the authors assessed the effect of olfactory imagery priming for participants classified as good or bad 'namers'. A key finding was an interaction where *d'* scores were selectively improved in the odour imagery priming condition (compared to visual imagery priming and a control condition), and this effect only occurred for the good 'namer' group. Importantly, the null effect of visual imagery priming suggests this facilitation was not a semantic effect, as a similar advantage to the olfactory condition would be expected. Instead, the authors suggest the good 'namers' have strong odour-name associations, which are reciprocally activated when attempting to imagine an odour. That is, only where an odour-name association is strong will an attempt to imagine an odour produce imagery priming effects. The strong odour-name association is proposed to reciprocally allow the activation of an odour image, and is similar to other demonstrations of a perceptual odour imagery which have shown improved ability after a learned link between odour and its name (Stevenson, Case, & Mahmut, 2007; Sugiyama, Ayabe-Kanamura, & Kikuchi, 2006).

Together, these findings are suggested to support imagery that includes a sensorial-type representation (Kosslyn, 2003). For example, in their review Rinck et al. (2009) describes this imagery as the consequence of an activated long-term representation (engaged by sniffing when attempting to image the odour), which is subsequently used to generate the sensorial representation. However, it has been argued that these findings can be accommodated as a capacity for odour imagery which is not available to consciousness (Stevenson, 2009; Stevenson & Attuquayefio, 2013). For example, Tomiczek and Stevenson (2009) propose that their findings supporting sensorial odour imagery may not be specific to a particular odour. Instead, they suggest that generic activation of olfactory neural networks (as a result of activating all extant odour-name associations) produces the priming effect without conscious imagery of a specific odour. It should also be noted, however, that such an interpretation contrasts the findings in Djordjevic et al., (2005), where an odour detection advantage was shown when the odour image and olfactory stimulus were matched, compared to when participants detected a different odour to the one they were required to image.

Internal representations that are unavailable to consciousness are also considered within the olfactory-centred unitary model described in Section 1.1.3, which suggests that demonstrations of olfactory memory is simply the result of residual activation and decay processes (Wilson & Stevenson, 2006). To be clear, it is suggested that olfaction can demonstrate phenomenal consciousness where there is experience of an olfactory sensation, but conscious access to these contents (through attention, or working memory) may not be possible (Stevenson, 2009). Consequently Stevenson (2009) suggests the effects of priming, hallucinations, and overlapping neural activations may be supported by unconscious imagery, but access to this internal representation for active maintenance or other tasks may not be possible.

However, it should be noted that there are other accounts of consciousness that suggest dissociation between phenomenal and access consciousness is not appropriate. Specifically, they suggest consciousness should instead be considered as a hierarchy of access to featural and semantic information (Kouider, de Gardelle, Sackur, & Dupoux, 2010). This alternative model describes levels on a hierarchy that are accessed independently, and thus allows a graded form of access consciousness. Graded access to consciousness may explain individual differences in olfactory imagery, and indeed a continuous scale of olfactory imagery ability has been proposed based on participant expertise (Arshamian & Larsson, 2014). Olfactory imagery is more vivid in olfactory experts than non-experts (Gilbert, Crouch, & Kemp, 1998), and expertise is associated with imagery that is consciously accessible (Plailly, Delon-Martin, & Royet, 2012; Royet et al., 2013; Stevenson & Attuquayefio, 2013). Olfactory imagery capacity is drastically affected by individual differences (Arshamian & Larsson, 2014), and may be related to semantic knowledge (Stevenson et al., 2007; Tomiczek & Stevenson, 2009) or perceptual experience (Delon-Martin, Plailly, Fonlupt, Veyrac, & Royet, 2013; Plailly et al., 2012). In Delon-Martin et al. (2013), for example, structural reorganisation of olfactory brain regions related to imagery was observed for those with extensive olfactory experience (perfumers). However, expertise is not necessarily essential in imagery. For example, other findings have shown an advantage for self-reported olfactory imagers in a same-different memory task was unrelated to the ability to identify odours (Köster et al., 2014). Taken together, there is some support for an ability to consciously access a perceptually-based olfactory representation, though expertise may be necessary for imagery and related working memory functions to occur that are analogous to other modalities.

1.2.4 Olfactory attention

Related to olfactory consciousness and working memory for odours is the role of attention in olfactory processing (White, 2012). For non-olfactory sensory systems, a close relationship between attention and working memory has been proposed (Lückmann, Jacobs, & Sack, 2014; Shinn-Cunningham, 2008). In particular, attentional focus is used to determine what enters memory (e.g. Broadbent, 1958), and maintenance of information in working memory can shift selective attention to those items (Awh, Jonides, & Reuter-Lorenz, 1998; Downing, 2000). Whether olfactory processing shares this two-way relationship, and might therefore be accommodated within a general model of attention (e.g. Chun, Golomb, & Turk-Browne, 2011; Knudsen, 2007), is equivocal. Other sensory systems will route information via the thalamus before projecting onto the cortex (see Guillery & Sherman, 2002). This region is thought to play an important role in directing selective attention (De Bourbon-Teles et al., 2014; Portas et al., 1998), and the route proposed to be responsible for the conscious perception of a stimulus (Pinault, 2004). However, olfactory processing also includes a direct route to the cortex that does not interact with the thalamus, which could indicate a unique relationship with selective attention and consciousness (Ongür & Price, 2000).

Chun (2011) describes two forms of attention, where external attention refers to the application of attention to sensory information, and internal attention refers to processes acting upon internally generated representations. Some aspects of olfactory external attention are unique. For example, Mahmut and Stevenson (2015) showed that an odour that has been habituated will not return to consciousness if an attempt is made to re-attend to that odour. That is, a person might habituate to the cooking smells within their house, and will be unable to consciously re-access this odour until the stimulus is removed and re-presented some time later (e.g. by leaving the room and coming back). In contrast, a repeated sound such as a ticking clock can be habituated, but easily re-attended if necessary (Mahmut & Stevenson, 2015).

The process of sniffing itself has a functional role in the allocation of external attention (e.g. Verhagen, Wesson, Netoff, White, & Wachowiak, 2007). For example, whilst olfactory attention does not allow spatial shifts of attention in the same way that visual attention can be used to select items from across the visual field, there is some ability to localise a monorhinally-activated nostril using an active sniffing process (Frasnelli, Charbonneau, Collignon, & Lepore, 2009). Olfactory attention to a particular temporal window can also be used to detect changes in the olfactory environment, and is mediated by the speed of a sniffing process (Mainland & Sobel, 2006; Verhagen et al., 2007). Furthermore, although it may not be possible to independently attend to multiple odours within a mixture (Jinks & Laing, 1999), participants have demonstrated an ability to direct attention to detect an odour quality in advance of its presentation (e.g. Gottfried & Dolan, 2003). That is, if a participant is looking at a picture of a fruit, the detection of that fruit's odour is faster than if looking at a picture of an unrelated item. This attentional priming is possible for both olfactory and visual information, reflecting similar processes across both domains (Gottfried & Dolan, 2003; Keller, 2011).

Post-perceptual processing, or internal attention, is related to rehearsal and refreshing of internal representations (Awh et al., 1998; Baddeley, 1986; Raye, Johnson, Mitchell, Greene, & Johnson, 2007), and in visual and verbal literature is also thought to be linked to the capacity for conscious imagery (Baddeley & Andrade, 2000). However, in olfaction the research on this is limited, and as previously discussed, the findings are equivocal on whether olfactory short-term memory includes the access to consciousness required for short-term maintenance processes (Stevenson, 2009; White, 2012; Zucco, 2003). Therefore, the next sections discuss the proposed capability for working memory in olfaction, with consideration of executive control, focussed attention, and maintenance. Examining these processes is important for understanding the position of olfactory memory ability in extant models of memory.

1.3 Olfactory working memory

Working memory describes the system used for temporary storage and manipulation of information, and which provides access to a stored representation required for goal-oriented behaviour (Baddeley & Hitch, 1974; Oberauer, 2009). This conceptualisation of working memory has been demonstrated through procedures requiring participants to maintain information whilst simultaneously performing processing tasks, such as the class of tests known as complex span (Daneman & Carpenter, 1980). Further tasks have also been developed, such as the n-back procedure, which are used to assess short-term maintenance of items whilst continually updating the items being rehearsed (e.g. Nystrom et al., 2000). However, these measures of working memory differ in their task requirements and in the possible executive functions involved (Redick & Lindsey, 2013; Schmiedek, Hildebrandt, Lövdén, Lindenberger, & Wilhelm, 2009; Wilhelm, Hildebrandt, & Oberauer, 2013; See Chapter 3 for a full discussion). Despite differences, however, measurements of working memory capacity have shown a strong link between working memory and higher-order abilities including language comprehension, problem solving, planning, reasoning, and intelligence (Conway, Kane, & Engle, 2003; Cowan, 2010; Oberauer, 2009; Süß, Oberauer, Wittman, Wilhelm, & Schulze, 2002).

This Introduction has presented evidence for a form of short-term memory that allows the temporary maintenance of olfactory stimuli (White, 1998, 2009), though whether this should be considered as a separate mechanism to long-term olfactory memory remains debatable (White, 2009; Wilson & Stevenson, 2006). Attempts to fit olfaction into a general attentional model (Knudsen, 2007; White, 2012) stipulate that not only is the short-term retention of odours necessary, however, but also that the information must be manipulated in line with the definition of working memory (Baddeley & Hitch, 1974; Cowan, 1999; Oberauer, 2009). To be clear, this means that odours must not only be remembered over short periods of time, but that they can be actively held on-line whilst goal-oriented functions are performed. The processes responsible for coordinating goal-oriented behaviour are known as executive functions, and are prominent components in models that describe cognitive control (Baddeley, 2012; Baddeley & Hitch, 1974; Cowan, 1999; Oberauer, 2009). Though specific interpretations differ, the executive processes are typically related to the control of attention and coordination of cognitive resources (Logie, 2011). Consequently, if olfactory working memory is similar to memory for other stimulus types, stored olfactory information must be available to the executive processes described in such models. For example, mental processes should be able to act upon stored odour representations to update them with new information (e.g. the n-back task), or features such as quality or intensity held on-line for multiple comparisons (e.g. discriminability tasks, or the triangle test described below).

Tentative support for olfactory working memory comes from odour detection within an odour-taste mixture (White, 2012). Using a triangle test and a 2-out-of-5 test, participants were required to identify the presence of benzaldehyde in a strawberry flavoured drink (that is, an odour-flavour mix) (Dacremont & Valentin, 2004). The triangle test involved single presentations (i.e. no re-tasting was allowed) of three strawberry drinks where a contaminated ‘odd one out’ must be identified. The 2 out of 5 task represented a more challenging version, where multiple comparisons were required to distinguish from 5 identical odours the two contaminated samples. The tasks are proposed to reflect manipulation in working memory because they require maintenance of previously stored odours (and tastes), and retrieval of each item for comparison to the item being evaluated. Importantly, participants with greater odour memory spans showed greater discrimination ability in these tasks, which is suggested to reflect greater working memory resources available for discriminating these odours (Dacremont & Valentin, 2004; White, 2012).

Stronger support for olfactory working memory capabilities is shown using the n-back task. This is described as a *maintenance + manipulation* task (Ragland et al., 2002), because it requires a rehearsal window to be constantly updated as new stimuli are presented (See Chapter 3 for a full discussion of the n-back task and strategies involved). Above-chance performance in the n-back task therefore support the ability to apply, amongst others, an updating executive function to the maintained set in working memory (Dade, Zatorre, Evans, & Jones-Gotman, 2001; Jönsson et al., 2011). Dade et al. (2001) demonstrated similar prefrontal activations (dorsolateral, ventrolateral, and frontal polar cortices) during both olfactory and facial stimuli n-back tasks, indicating similar engagement of working memory processes regardless of stimulus modality. However, the study is criticised for its use of familiar odours, meaning it is unclear the effects verbal labelling may have had on the working memory resources employed (Jönsson et al., 2011). To be clear, if participants are relying upon verbal labels for the odours, performance is instead demonstrable of verbal rather than olfactory working memory. This was addressed in Jönsson et al. (2011), who showed above-chance performance in the n-back task for odours that were difficult to verbally label, thus supporting working memory updating for odours in the absence of verbal recoding (though performance was improved with odorant verbalisability).

In summary, there is above-chance performance in an olfactory task that requires manipulation of the representations though continual updating (Dade et al., 2001; Jönsson et al., 2011), and an ability to focus attention on stored odour representations in order to make comparisons for effective stimulus identification (Dacremont & Valentin, 2004). Together, this (albeit limited) evidence lends support for olfactory working memory, though some questions remain regarding the importance of verbal processing for effective application of working memory processes.

1.3.1 Maintenance of information: Rehearsal and Refreshing

A question in olfaction is how this information might be held active in memory over a delay. For other modalities, maintenance of information is proposed to arise from *rehearsal* (one or more items cycled in a loop multiple times) or *refreshing* (an instance of reflective attention) (Cowan, 1992; M. K. Johnson et al., 2005). Refreshing is a basic executive process whereby a recently-activated representation is re-attended in order to maintain its memory trace in an active state (Barrouillet, Bernardin, & Camos, 2004; M. K. Johnson, 1992).

Verbal information is strongly associated with the use of rehearsal (that is, the overt or covert articulation of phonological codes) to maintain information (Baddeley, Lewis, & Vallar, 1984), and is supported through impairments to memory as a result of word length (Baddeley, Thomson, & Buchanan, 1975) or similarity (Baddeley, 1966). However, maintenance of information can also benefit from *attentional refreshing*, which unlike verbal rehearsal is not specific to phonological codes (Camos, Mora, & Oberauer, 2011). Refreshing may be performed in addition, and separately, to articulatory rehearsal (Camos, Lagner, & Barrouillet, 2009; Hudjetz & Oberauer, 2007). This refreshing process is proposed to be a general-purpose attentional-maintenance mechanism, which can be supplemented by a specialised phonological rehearsal process when this information is available (Camos et al., 2009). The two processes differ both temporally and in the amount of items retained, where rehearsal allows retention of multiple items over several seconds, whilst refreshing increases the activation of a memory trace only momentarily (Raye et al., 2007). In Raye et al., (2002), greater activations in the left dorsolateral prefrontal cortex were observed when requiring participants to think back to a presented item when cued (refreshing), compared to when a target is re-presented visually to be vocalised (rehearsal), and to other cued tasks such as to think ‘dot’ when presented with a dot, or to think a particular direction when presented the appropriate.

The use of rehearsal or refreshing for non-verbal information is less clear, however. Though an analogous system of rehearsal has been proposed for visuo-spatial information (e.g. Logie, 1995) which may be separate to refreshing (see Raye et al., 2007 for a discussion), it is unclear how the two may differ. Indeed, most explanations of non-verbal maintenance now simply describe refreshing, which has been proposed as a potential maintenance mechanism for information including pictures, textures, and words, and also for maintaining bound representations in the episodic buffer (Baddeley, 2012; M. K. Johnson et al., 2005; M. R. Johnson, McCarthy, Muller, Brudner, & Johnson, 2015). Evidence also supports that this refreshing process improves the accessibility of representations in working memory (Souza, Rerko, & Oberauer, 2015). Consequently, refreshing may be an appropriate executive function for processing in working memory when an articulatory rehearsal mechanism is not available.

Importantly, refreshing is described as both a maintenance and manipulation process. An example of this might be allowing a representation to be strengthened relative to others in a sequence (M. Johnson et al., 2005). Consequently, refreshing is suggested as an important process in the n-back task where items in a maintenance window must be recollected in their correct serial position (M. K. Johnson, Raye, Mitchell, Greene, & Anderson, 2003; M. R. Johnson et al., 2015; Raye, Johnson, Mitchell, Reeder, & Greene, 2002). Using refreshing, matching a target to the *n*th item can be made possible by increased activation strength of that item relative the other items in the rehearsal window (see Juvina & Taatgen, 2007 for an outline of n-back control strategies). Activation in the left PFC during the n-back task has been associated with refreshing, and activation in the right PFC has been associated with target matching (Cohen et al., 1997).

If some form of active maintenance process is possible for olfactory representations, it is presumably the latter non-verbal refreshing process that is responsible. However, the possible contribution of verbal information to the representation of olfactory information (e.g. Jönsson et al., 2011) means a rehearsal process should also be considered in the performance of olfactory working memory tasks. It should be noted that evidence for olfactory n-back updating (Dade et al., 2001; Jönsson et al., 2011) also suggests that rehearsal or refreshing of olfactory information is possible, though it is debated in Sections 1.2.3 whether an internal olfactory representation is available to consciousness (see Djordjevic et al., 2005; Plailly et al., 2012; Stevenson, 2009; Tomiczek & Stevenson, 2009). Neuroimaging evidence supports a relationship between refreshing and imagery, where selective regions associated to the modality being refreshed are activated (in addition to activation in the dorsolateral prefrontal cortex which is associated to the domain-general top-down allocation of attention, Curtis & D'Esposito, 2003; M. R. Johnson et al., 2007). That is, the modality-specific activations reflect regions typically activated in the presence of an actual stimulus, and suggest the activation of an internal representation, or *image,* of the stimulus during refreshing (Ranganath & D'Esposito, 2005).

Though there is limited research specifically concerning rehearsal in the olfactory modality, the unique characteristics of olfactory memory present conflicting evidence on whether rehearsal or refreshing strategies can be employed. For example, the inclusion of a same-modality item between learning and test will typically interfere with maintenance of items, but in olfaction an effect of retroactive interference is not always observed (Lawless & Engen, 1977; Schab, 1991; Zucco, 2003). Absent retroactive interference is evidence against a rehearsal or refreshing process, as it suggests that additional items are not interfering with maintenance of the target memoranda (Stevenson, 2009). Indeed, Stevenson (2009) argue that the instances where retroactive interference has been observed in olfactory memory may be attributed to perceptual similarity (discussed in Section 1.2.1, e.g. Andrade & Donaldson, 2007; Walk & Johns, 1984). Furthermore, as discussed in Section 1.2.2, primacy effects may be related to rehearsal of items early in a list (cf. Tan & Ward, 2007), and have been absent in serial position tasks that typically demonstrate primacy for other modalities (A. J. Johnson et al., 2013; A. J. Johnson & Miles, 2009; cf. Reed, 2000).

In summary, the evidence for a refreshing process in olfactory memory is equivocal, and is likely dependent on the ability to form a consciously accessible internal image (see Sections 1.2.3 and 1.2.4). Arguments against refreshing are evidenced by the absence of retroactive interference (Zucco, 2003), and the lack of a primacy function in serial memory (e.g. A. J. Johnson & Miles, 2009), though other research has shown both features in olfactory memory (e.g. Andrade & Donaldson, 2007; Reed, 2000). Above, the ability for external and internal attention to odours, and evidence for the use of odours in an n-back task (Jönsson et al., 2011), supports a working memory ability that can process olfactory information. In Section 1.4, such an ability is discussed with consideration of models of working memory and executive functioning.

1.3.2 Executive functions

An important issue identified in this Introduction is whether controlled working memory resources can act upon a stored olfactory representation in memory. That is, whilst recognition memory of a perceptually represented odour is supported (Andrade & Donaldson, 2007; Møller et al., 2004; White et al., 1998; Zelano et al., 2009), the allocation of attention to this representation, and whether mental operations can be performed on them, has received far less scrutiny. In working memory, these operations are called executive processes, and are proposed to be the modality independent application of attention required to complete a task (e.g. Baddeley & Hitch, 1974; Miyake et al., 2000). To be clear, although there is debate over whether executive functions themselves should be fractionated, they are typically unrelated to the modularity debate offered in Section 1.4. This section gives a brief outline of the nature of executive functioning, and the processes that may be engaged during demonstrations of olfactory working memory capacity.

Executive functioning has received considerable debate in the literature, with particular contention over whether it should be considered not as a single executive resource for controlled attention, but as a fractured system that consists of multiple functions. Low correlations between tasks that ostensibly measure distinct executive functions support fractionation of the executive (Lehto, 1996), though this may be attributable to the *task-purity* problem where different (non-executive) processing requirements mask the presence of a common executive ability (Miyake et al., 2000; Miyake & Shah, 1999). However, influential in this area is the identification of three executive function latent variables, of *updating*, *inhibition of prepotent responses*, and *task-set shifting* (Miyake et al., 2000). The updating process involves maintenance of items and replacement of no longer relevant representations in declarative working memory, and is likely an important process in working memory tasks such as the n-back procedure (Oberauer, 2009). Inhibition refers to the deliberate suppression of prepotent responses such as those generated in the stroop task, whereas task-set shifting concerns the ability to switch between multiple tasks, operations, and mental sets (Miyake et al., 2000; Oberauer, 2009). Analysis of latent variables provides strong support for separate processes, as it minimises the task-purity issue by extracting the common variance in each individual task (Miyake et al., 2000). Furthermore, assessing how these latent variables predict higher-order cognition may support distinctive executive functions. For example, updating tasks have been shown to be closely related to measures of intelligence, whilst the other two functions have shown a much weaker relationship (Friedman et al., 2006).

The constructs described in a fractured executive were moderately correlated, however, and it has been argued that this is because they contain a common mechanism (Engle, Tuholski, Laughlin, & Conway, 1999; Friedman et al., 2008; Jurado & Rosselli, 2007; Mccabe, Roediger III, Mcdaniel, Balota, & Hambrick, 2010; Wilhelm et al., 2013). Typically these arguments retain the dissociable components of executive functioning (e.g. Miyake et al., 2000), but suggest there is a general attentional function involved that is not related to general intelligence or perceptual speed (Banich, 2009; Friedman et al., 2008; Garon, Bryson, & Smith, 2008). Furthermore, Mccabe et al. (2010) showed an underlying component in constructs representing both executive function tasks and working memory tasks, which they called executive attention. This cognitive ability is proposed to reflect focused attention, which is necessary during goal-oriented activity, and is present in tasks that ostensibly tap into executive functions and also in tests of working memory capacity (Mccabe et al., 2010). Alternatively, Oberauer (2009) suggests the primary process in working memory is a specific binding process. That is, in a working memory task, an item is bound to its context, and it is this bound representation that can be retrieved and updated. They propose this as the reason why updating tasks are strongly related to working memory capacity, because they measure the ability to quickly retrieve and update these bindings (Oberauer, Süß, Schulze, Wilhelm, & Wittmann, 2000).

In summary, models of executive function generally support its fractionation into specific functions such as those described in Miyake et al. (2000). However, there is some discussion over whether tasks that appear to measure these functions, and also tests of working memory capacity such as complex span, should be considered in terms of their use of high-level attentional mechanism (e.g. Mccabe et al., 2010). Clearly important in these functions is the ability to focus attention on the stimuli and task-goals, and inhibit attention where necessary (Jurado & Rosselli, 2007). Regardless, the separation of updating, inhibition of prepotent responses, and task-set shifting, is clearly a useful taxonomy for understanding the role of attention in different tasks. The present thesis therefore not only explores temporary storage of olfactory information, but assesses olfactory working memory using an n-back task, a procedure proposed to require an updating process (Wilhelm et al., 2013).

1.4 Models of working memory

This section considers competing theories surrounding the structure of working memory. These are typically categorized as modular accounts, where independent modules exist with specific functions with their own limited capacity, and unitary accounts, which propose a single system specialised for processing activated information from long-term memory. An in-depth assessment of evidence for the multicomponent working memory framework (Baddeley, 2000; Baddeley & Hitch, 1974) is presented, and the procedures used for demonstrating separation of each process are evaluated. Furthermore, evidence for a distinct olfactory module in this framework is discussed. In Section 1.4.2, the embedded-processes unitary model is described (Cowan, 1999), and is discussed with consideration of the olfactory-centred unitary model.

1.4.1 Modular accounts

Some theorists propose the underlying mechanisms of memory to be *modular*, where multiple systems operate independently from each other. The most well-known, though by no means the only, modular account of working memory is the multicomponent working memory framework. This outlines a limited capacity system that makes use of multiple storage and processing resources that act in concert (Baddeley, 1986; Baddeley & Hitch, 1974). This includes modality-specific systems for storage of information; specifically, the phonological loop for verbal storage and the visuo-spatial sketchpad for storing visual and spatial information (Baddeley, 1986; Baddeley & Hitch, 1974; Logie, 1995). These slave systems are considered distinct from domain-general control and allocation of attentional resources, which is regulated by a central executive (Baddeley, 1986). That is, a central executive is proposed to focus and divide attention when necessary, is able to switch attention between tasks, and can interface with long-term memory (Baddeley, 2012).

Utilisation of these modality-specific stores is not mutually exclusive; both stores can be employed to maintain the same item. For example, when viewing a visual image, an individual may store a verbal description of that image in addition to the iconic representation (Logie, Della Sala, Wynn, Baddeley, & Sala, 2000). Consequently, a third slave system, the episodic buffer, has been incorporated into the model to allow working memory resources to act upon bound multidimensional representations (Baddeley, 2000, 2012). Though originally proposed to require attentional focus (Baddeley, 2000), the process of binding information from within a slave system is now thought to be relatively automatic with the buffer itself acting as a passive store (Baddeley, Allen, & Hitch, 2011). However, binding of verbal and visual features is disrupted when a concurrent task is performed, suggesting differences in attentional requirements when the bound item contains information from across slave systems (Allen et al., 2012; Elsley & Parmentier, 2009). Stored items within the episodic buffer may represent multiple forms of binding; for example, temporary bindings are required in order to perform several working memory tasks, whereas durable bindings may occur when new information is attached to a context in long-term memory (Baddeley, 2012). Furthermore, the buffer is assumed to provide access to conscious awareness, though whether this means people are not aware of content of other subsystems until it is passed to the buffer is unclear (Baddeley, 2012). Finally, Baddeley (2012) speculates that the episodic buffer is a suitable location for refreshing-based rehearsal of stimulus types other than verbal and visuo-spatial stimuli, and that, for example, smell and taste information may be fed into this buffer from their own subsystems.

The multicomponent model proposes two independent sub-systems for processing phonological and visuo-spatial information (Baddeley, 1986; Baddeley & Hitch, 1974; Logie, 1995). Processing verbal information within the phonological loop relies on storage within a phonological store, and constant refreshing of memoranda using vocal or subvocal rehearsal (Baddeley et al., 1984). Storage as a phonological code is supported by poorer memory performance for phonologically similar items (i.e. the phonological similarity effect (PSE); Conrad, 1964; Salamé & Baddeley, 1986), and the process of rehearsal evidenced by reduced span for items that take longer to articulate (i.e. the world length effect (WLE); Baddeley et al., 1975; Mueller, Seymour, Kieras, & Meyer, 2003). Whilst auditory-verbal stimuli have direct access to the phonological store, visual-verbal stimuli require phonological recoding within the phonological loop. Support for this is found with abolition of both the WLE (Baddeley et al., 1975), and PSE (Larsen & Baddeley, 2003; cf. Longoni, Richardson, & Aiello, 1993 for verbal auditory information; Saito, Logie, Morita, & Law, 2008) for visual-verbal (but not auditory verbal) stimuli following concurrent articulation (CA). This detrimental effect of irrelevant articulation, during a procedure that is thought to occupy the articulatory loop, provides strong evidence for the use of a sub-vocal phonological rehearsal process.

The use of CA is an example of *dual-tasking*, and provides compelling evidence for independent verbal and visuo-spatial systems. To be clear, strong evidence for modularity is found when concurrent different modality memory tasks produce little deficit to performance, compared to a large performance drop when these tasks operate the same modality (Baddeley, 1986; Baddeley & Hitch, 1974). Since this paradigm is considered compelling evidence for modality-specific slave systems, and is later considered in the context of evidence for an olfactory buffer (Andrade & Donaldson, 2007), evidence from dual-tasking studies is considered in detail below. In early studies, Baddeley and colleagues applied CA (what they called articulatory suppression) to a free recall task by having participants speak a repeated word throughout the task, which was shown to dramatically decrease accuracy (e.g. J. T E Richardson & Baddeley, 1975). Whilst CA provides an example of interference with an articulatory rehearsal process, other secondary tasks may be performed where the modality and level of interference is adjusted. For example, Meiser and Klauer (1999) performed an in-depth analysis of dual-task effects on sequence memory for multiple modalities. Across 6 experiments, secondary tasks required concurrent articulation or tapping, designed to differentiate between same-modality (i.e. slave-system interference) and cross-modality (i.e. executive function interference) effects. Furthermore, these tasks were performed during either encoding or retention, and engaged lower loads on central executive resources as *solid-state* (vocally repeat a single letter or tap a single button), or high loads as *changing-state* (repeat an alphabetical sequence of letters or tap in a clockwise direction) tasks. Together, the manipulations allowed assessment of tasks that differ across executive resources, performed across differing stages of a working memory task, and upon the same or different modalities. They demonstrated support for the multicomponent working memory framework by showing a dissociation of concurrent task effects performed in the retention interval, where CA impaired verbal sequence memory more than spatial sequence memory, and tapping impaired spatial more than verbal sequence memory. In addition, greater interference from the high-load *changing-state* concurrent tasks was observed only during the encoding stages, which suggests an important role of executive resources when encoding item or serial position information.

Similarly, other dual-task studies have used complex secondary tasks to load memory or executive resources. Cocchini et al. (2002) had participants perform either a digit sequence or pattern memory task, during the retention interval this was paired with either the alternative memory task, a perceptuomotor tracking task (Experiment 1), or CA (Experiment 2). Their findings in both experiments showed almost absent interference when memory tasks occupied different domains, and this was taken as support for processing independence of digit and pattern information. The tracking task did not impair digit memory, and only minimally affected pattern memory. Furthermore, there was substantial interference from CA on only the digit memory task, and this was interpreted as evidence for rehearsal in the phonological loop. Together, the results support a multicomponent model interpretation where domain-specific slave systems are utilised in working memory.

The above evidence demonstrates that interference is typically far greater if two tasks are performed from within the same domain, compared to when tasks are performed from separate modalities (Cocchini et al., 2002; Meiser & Klauer, 1999). Cross-modal interference however, though often much smaller than the interference from within-modality dual tasks, is non-trivial in several experimental studies that claim a double dissociation (Jones, Farrand, Stuart, & Morris, 1995; discussion in Morey, Morey, van der Reijden, & Holweg, 2013). This effect may be explained within a multicomponent framework as occurring due to the general cost on the domain-general central executive from performing two tasks simultaneously (Logie, 2011), or by the recoding of items in another modality (Paivio, 1990). However, asymmetric cross-modal interference has also been found, where verbal working memory is more robust to interference from visuospatial tasks than visuospatial working memory is to interference from verbal tasks (C. C. Morey et al., 2013). Morey and colleagues suggest these findings may be better explained by a domain-general storage and attention processes, such as those suggested by Cowan et al. (2005) or Oberauer (2009), but with additional verbal-specific processes to account for the resilience of verbal memory to interference. That is, some verbal-specific store or rehearsal mechanism may need to be accommodated within these models, or a specialised process that supports verbalisation which is not available for visualisation (C. C. Morey et al., 2013).

Dual-task studies have informed other aspects of working memory; for example, interference is also observed when both tasks contain an order component regardless of the modality of these tasks (Depoorter & Vandierendonck, 2009; Vandierendonck, 2016). Interference for two different modality order memory tasks would not be expected if the encoding of order information takes place within modality specific subsystems. In contrast, these findings instead suggest a modality-independent system for serial recall. Vandierendonck (2016) consider this finding incompatible with a multicomponent model, unless the model is adapted to allow an item-position binding to be maintained in the episodic buffer, or the central executive. However, support for modality-specific serial memory components is shown by sequence memory for visually presented verbal material exhibiting both phonological similarity effects (indicating storage of order information as phonological codes), and visual similarity effects (indicating the additional use of visual codes) (Saito et al., 2008; see also Guérard & Tremblay, 2008 for double dissociation interference effects in serial order memory). Effects from both types of similarity for the same materials suggest an independent contribution of visual codes in the retention of these verbal sequences, and this is also supported by an effect of concurrent articulation in abolishing the PSE only (Saito et al., 2008). Saito et al. (2008) conclude that domain specificity is necessary in models of order memory because there is a clear ability to retain serial order with visual-verbal stimuli despite suppression of phonological codes. However, it should be noted that an amodal mechanism for processing serial order could also explain these findings, by modality-specific item memory and a modality-nonspecific memory for order. Indeed, the use of these visual codes in order memory relies on similar principles to the use of phonological codes. This is observed from the similar effects of visual similarity and phonological similarity on sequence memory (Saito et al., 2008), and also from similar serial-position functions in visual memory even when phonological coding is discouraged (Avons, 1998; Hurlstone, Hitch, & Baddeley, 2014; A. J. Johnson & Miles, 2009; Logie, Saito, Morita, Varma, & Norris, 2016).

In summary, dual-task studies may load the limited resources available in working memory, and are useful for studying the separation of cognitive processes. Though double-dissociations are compelling evidence for modular systems in working memory, cross-modal interference effects suggests a secondary task may simply load upon the executive resources available and this should be considered in any dual-task interpretation (e.g. C. C. Morey et al., 2013). Relatedly, loading executive resources has been proposed to disrupt cross-modal binding processes (Allen et al., 2012; Elsley & Parmentier, 2009), which should be considered when assessing the influence of secondary tasks on working memory task performance.

Neuropsychological evidence provides a further source of evidence to assess modularity in working memory. Vallar and Baddeley (1984) provided early evidence for selective impairment to phonological working memory. Patient PV suffered a left-hemisphere stroke, after which she showed impairments in immediate memory when stimuli were presented in the auditory modality, and did not show a phonological similarity effect when items were presented visually. Importantly, these effects occurred despite no articulatory problems, suggesting the impairments were localised to the phonological store. In addition, Hanley, Young, and Pearson (1991) presented a case study of ELD, who had suffered a right-hemisphere aneurysm. The pattern of impairments were localised to visuo-spatial memory (Brooks Matrix and Corsi Blocks), whilst verbal sequence memory was intact. Together, these findings support a double dissociation where the phonological and visuo-spatial slave systems can be selectively impaired (see Baddeley, 2007 for a further discussion of neuropsychological evidence).

How olfactory information may be accommodated within the above modular conceptualisation of memory is currently unclear, though is speculated by Baddeley (2012) to be processed within the episodic buffer. However, Baddeley does not rule out input from an olfactory slave system, and research has supported such an olfactory-specific subsystem in working memory (Andrade & Donaldson, 2007; Zelano et al., 2009). An independent olfactory subsystem may provide a suitable location for the pattern-matching process described in Stevenson and Boakes (2003), and explain the qualitative differences historically observed for olfactory memory (e.g. Engen et al., 1973). Furthermore, like other modalities, information initially processed within the olfactory buffer can interact with phonological information, visuo-spatial information, and long-term semantic memory (Baddeley, 2000). Bound representations that include perceptual and semantic features, stored in the episodic buffer, may by consistent with the features described in the object-processing account of olfactory memory (Wilson & Stevenson, 2006), or accounts that suggest a dual-representation for all odours (Lyman & McDaniel, 1986; Yeshurun et al., 2008).

Andrade and Donaldson (2007) examined evidence for the inclusion of an independent olfactory working memory subsystem within the multicomponent working memory framework (Baddeley & Hitch, 1974). They employed the classical method used for supporting modularity, a dual-task paradigm, and found that in Experiment 1, a primary verbal task was not affected by a secondary olfactory task (and to the same extent as a secondary visual task). This was used as evidence against the proposition that olfactory memory simply reflected rehearsal of verbal labels for the odours; as this would have necessitated utilisation of the phonological store for both the verbal and olfactory tasks. However, despite such a prediction, interference was not found. Furthermore, in Experiment 2, a primary olfactory recognition task was impaired to a greater extent by a concurrent olfactory task than a concurrent verbal or visual task. To be clear, if remembering odours was a de facto verbal task, a secondary olfactory task should be as detrimental to primary olfactory task as a secondary verbal task. This was not the case. Taken together, the findings support an olfactory memory subsystem which makes use of specialised, independent resources in working memory. That is, they suggest that verbal processing is not an intrinsic part of olfactory memory, despite facilitative effects of verbal labelling (Andrade & Donaldson, 2007).

Support for this independent system has also been observed in neuroimaging research. In a delayed-match-to-sample working memory task, sustained activity in the inferior frontal gyrus was demonstrated during memory for nameable odours (consistent with information rehearsal within the phonological loop) (Zelano et al., 2009). In contrast, when using hard-to-name odours, the same working memory task resulted in activation within the piriform cortex. This activation pattern indicates a dedicated mechanism for olfactory processing, which can be utilised in the absence of verbal identification for odours. Such dissociable neurological activations also suggest that nameable and hard-to-name odours might be processed in qualitatively different ways. However, there was, importantly, both residual activation in the piriform cortex for nameable odours, and residual activation in the inferior frontal cortex for hard-to-name odours that was greater than the activations observed for the same task using auditory stimuli. This suggest that differential brain activation may be weighted based upon the extent to which an odour is verbalisable. However, whilst Zelano et al. (2009) argue that both working memory stores (verbal and olfactory) are utilised to some extent in maintaining an odour image, they suggest that the minimal use of verbal processing for an unnamed odour reflects a general categorisation label (e.g. "nice" / "nasty" etc.). Zelano et al. (2009) also note other findings where activation in the olfactory cortex will increase following presentation of the name of an odour (González et al., 2006), and also from a visually-presented object that is related to a smell (Gottfried & Dolan, 2003). This cross-modal effect is similar to the ability to form a visual image from a verbally described object, or a phonological representation from a visually presented word, which they suggest is evidence for an olfactory *flacon* (buffer) that is comparable to the phonological loop and visuospatial sketchpad.

A dual-representation of perceptual and verbal information during short-term memory for odours, even when hard-to-name, is also proposed by Yeshurun and colleagues. Early olfactory processing is proposed to be ipsilateral in nature, meaning activations at the level of olfactory receptors and the olfactory bulb do not cross hemispheres (e.g. Lascano, Hummel, Lacroix, Landis, & Michel, 2010). Using this knowledge, they presented a model of olfactory working memory, based on behavioural effects using monorhinal presentation of nameable and hard-to-name odours, and a same-nostril/different-nostril manipulation across target and probe presentation (Yeshurun et al., 2008) To be clear, any performance changes when target and probe were presented across nostrils (compared to both presented to the same nostril) can be used to assess the level of perceptual processing used for performing the task. Indeed, the authors observed a nameable odour recognition advantage that was enhanced when target and probe were presented to different nostrils, and this pattern of results was used to falsify a perceptual-only or verbal-only representation. Specifically, Yeshurun et al. rejected several possible representations in memory for nameable and unnameable odours, (1) a *low-level perceptual-only* representation, (2) a *high-level perceptual-only* representation, (3) a *verbal* representation generated for only nameable odours, or (4) a *verbal-only* representation for all odours. These propositions are discussed in more detail below.

A low-level perceptual-only representation suggests comparisons occur in the olfactory bulb (involved in the early stages of establishing an olfactory pattern, Gottfried, 2010), which would predict improved recognition performance when target and probe were presented to the same nostril. This is because there is minimal exchange between the two bulbs, so matching the representations of target and probe at this level across nostrils would be impaired. In comparison, a high-level perceptual-only representation suggests processing in the piriform cortex (linked to hard-to-name odour processing in Zelano et al., 2009, and a possible location for the odour-object store in Wilson & Stevenson, 2006). However, perceptual-only processing in the piriform cortex would predict nameable and hard-to-name odours to be represented in one of two ways. First, the perceptual representation may be similar for both category of odour, and would predict similar performance when target and probe were presented across-nostrils. Alternatively, the perceptual representation may be more accessible for nameable odours, and would predict better nameable-odour performance when target and probe were presented to the same nostril. That a nameable advantage was observed, but this was only across nostrils, rules out a perceptual-only representation based on the logic above. Instead, some influence of verbal processing appears to be involved in a representation to explain the advantage for nameable odours. A role of verbal processes for only nameable odours was rejected, however. This is because given that odour processing is (primarily) ipsilateral, one might predict a left nostril advantage if these nameable odours are represented verbally, due to the left hemisphere's role in language. Again, this result was not found. Furthermore, a verbal-only representation was also rejected because of a lack of a general left-nostril processing advantage.

The authors therefore suggest the data fit a model where verbal and perceptual processes are utilised in a dual-representation for all odours. That is, there is clearly some use of a pure perceptual representation in working memory, and this might reflect the presence of an independent olfactory buffer. However, there is also an interaction with verbal information that may provide an additional cue for retrieval (Yeshurun et al. 2008). These findings may also, however, be accommodated within the olfactory-centred unitary model, where hard-to-name odours are suggested to rely on a low-level pattern matching system and broad verbal labels are prone to errors. In contrast, named odours use the same perceptual pattern that is linked strongly to a centrally-mediated representation reflecting an identified odour object (see R. A. Frank et al., 2011; Wilson & Stevenson, 2006).

In summary, the evidence discussed indicates the existence of an independent olfactory memory system within a modular framework that supports recognition memory, but naming will provide an additional means to facilitate the olfactory representation (Andrade & Donaldson, 2007; Stevenson & Mahmut, 2013a). Though there is evidence against a default influence of verbal coding (Andrade & Donaldson, 2007), other findings suggest that verbal information may be utilised to varying levels for all odorants (Yeshurun et al., 2008; Zelano et al., 2009).

1.4.2 Unitary accounts

Alternative to modular accounts are models that take an amodal and unitary approach to memory. The embedded-processes model (Cowan, 1999) is unitary to such extent that processes of short-term and working memory are the result of activated long-term memory representations (e.g. Cowan, 1999). Working memory is proposed to engage items that fall under a focus of attention embedded within a field of activated memory which includes sensory, phonological, and semantic features from across all modalities (Cowan, 2010). Rather than separating modalities into independent structures in working memory (e.g. Baddeley & Hitch, 1974), processing of verbal information, for example, can therefore simply be considered as just one of several forms of activated memory (Cowan, 2008).

The field of activated representations from long-term memory are proposed to reflect temporarily accessible items (i.e. short-term memory), which may or may not be available to consciousness (Cowan, 2010). Consequently, items placed in the focus of attention are subject to processes typically described as *working memory*, which is a capacity limited process but can be improved by combining items to form *chunks* (Cowan, 2001). These items can be refreshed to maintain their activation, will be processed to a greater depth than other items, and will be kept in mind to assist in working memory task (Cowan, 2010). It should also be noted that variations to this model exist. For example, Oberauer (2009) suggests a similar model, but described the multiple items in Cowan's focus of attention instead as a lower level of activation, and added a single-item focus of attention for processing in working memory.

The olfactory-centred unitary model described in Section 1.1.3 offers a detailed account of the perceptual processes that can lead to a perceptual representation of an odour in memory. In their model, Wilson and Stevenson (2006) suggest that the perceptual representation of an odour can be activated similarly to how other modalities in Cowan's embedded processes model are activated (Cowan, 1999). That is, the activation of a stored olfactory representation during odour perception is remembered by its residual activation in long-term memory (Wilson & Stevenson, 2006). However, these olfactory representations are proposed to be unavailable to consciousness (Stevenson, 2009). That is, these items may be inaccessible to the focus of attention that engages items for processing in working memory, particularly if semantic information is unavailable (Tomiczek & Stevenson, 2009; cf. Jönsson et al. 2011).

1.5 Summary

The above discussion outlines the equivocal findings related to olfactory working memory and conscious imagery. Despite the relative paucity of evidence there is some support for the ability to represent odours in short term memory (see White et al., 2015). Whether short-term memory for odours is similar to memory for other modalities is, as yet, unclear. For example, odours may be recognised through residual activation of odour engrams that are not available to consciousness (Stevenson, 2009; Wilson & Stevenson, 2006). Whilst evidence does suggest consciously accessing an olfactory representation is difficult, there is support for such an ability (Arshamian & Larsson, 2014; Arshamian, Olofsson, Jönsson, & Larsson, 2008; Tomiczek & Stevenson, 2009).

An attempt to fit olfaction into a general attentional model is reasonably well-supported by short-term memory evidence (White, 2012), in particular supporting the application of internal attention for rehearsal and working memory (Dade et al., 2001; Jönsson et al., 2011; Yeshurun et al., 2008; Zelano et al., 2009). The extent to which this is underpinned by verbal representations remains under debate, but there is some support for a perceptual representation in memory (White et al., 1998; Yeshurun et al., 2008) despite a proposed role of verbal information in most tasks involving explicit odour memory (Yeshurun et al., 2008; Zelano et al., 2009).

Chapter 2: Odorant normative data for use in olfactory memory experiments: Dimension selection and analysis of individual differences

2 Chapter Summary

The first aim of this thesis was to obtain a large normative database of commercially available odour stimuli. These normative data were assessed for their utility in the control and manipulation of odour stimulus characteristics. For example, the introduction to this thesis describes an important role of odour knowledge on olfactory perception and memory. It is important, however, to consider the variance from individual differences that may reduce the effectiveness of stimulus control based on normative information. The study below outlines the process of collecting these data, and its subsequent validation as suitable for use in stimulus control.

This chapter reports normative ratings for 200 food and non-food odours. One hundred participants rated odours across measures of verbalisability, perceived descriptive ability, context availability, pleasantness, irritability, intensity, familiarity, frequency, age of acquisition, and complexity. Analysis of the agreement between raters revealed that four dimensions, those of familiarity, intensity, pleasantness, and irritability, have the strongest utility as normative data. The ratings for the remaining dimensions exhibited reduced discriminability across the odour set and should therefore be used with caution. Indeed, these dimensions showed a larger difference between individuals in the ratings of the odours. Familiarity was shown to be related to pleasantness, and a non-linear relationship between pleasantness and intensity was observed which reflects greater intensity for odours that elicit a strong hedonic response. The suitability of these data for use in future olfactory study is considered, and effective implementation of the data for controlling stimuli is discussed.

2.1 Creation of a normative database

2.1.1 Introduction

Cross-modal comparison between olfactory memory and memory for other sensory modalities has produced mixed findings. Some studies have reported a pattern of memory consistent with other stimulus types (e.g. A. J. Johnson & Miles, 2007; Miles & Hodder, 2005; White & Treisman, 1997), whereas others have reported qualitatively different trends for olfactory stimuli (e.g. A. J. Johnson et al., 2013; A. J. Johnson & Miles, 2009; Reed, 2000). One possible interpretation of the latter finding is that olfactory memory differs qualitatively to that for other stimulus types and potentially resides within a separate olfactory-specific memory store (Andrade & Donaldson, 2007; Zelano et al., 2009).

An alternative explanation for the above disparity may relate to the criteria employed for odour selection. The characteristics of an odour can be an important determinant of memory performance, both quantitatively and qualitatively. Importantly, short of an a priori assessment of name-ability, there is limited control on the psychological characteristics of the odours. These odour characteristics may be of importance in determining cross-modal serial position function congruence, since the psychological distinctiveness of items (a somewhat ill-defined construct that can be influenced by perceptual familiarity) is argued to affect both the primacy and recency components of the serial position curve (Hay, Smyth, Hitch, & Horton, 2007).

One method by which the perceptual experience of odours can be assessed is from ratings of the odours across various dimensions. Judgments of this nature are typically obtained via subjective ratings pre-test (Yeshurun et al., 2008), during encoding (Larsson, Nilsson, Olofsson, & Nordin, 2004), or after the experiment through post-hoc data collection (M. J. Olsson et al., 2009). Indeed, there is some merit to collecting data this way, the most notable being mitigation of individual differences. For example, individual naming ability can allow tailored selection of odorants for use in subsequent memory and discrimination tasks (Rabin, 1988; Rabin & Cain, 1984). However, issues arise when tasks require novel presentation, and speeded encoding or recognition. In addition, these methods of odour stimuli categorization are often inconsistent, utilizing different scales and tasks, and resulting in these data rarely being used beyond the confines of the study in which they were collected. To this extent, the data are study-dependent. It is, therefore, desirable to have a reliable catalogue of odours and normative data which will facilitate the use of odours in olfactory memory research. Accordingly, the present study attempts to provide data norms for a large set of commercially available odours, analogous to that produced for words (Coltheart, 1981), faces (Ebner, Riediger, & Lindenberger, 2010), and objects (Yoon et al., 2004). Normative data in the verbal processing literature allows strict control of the orthographic, phonological, and psychological characteristics of words. An odour data analogue will thereby enable researchers to both strictly control for, and manipulate, levels of psychological difference.

There is some limited precedence for the use of normative data for olfactory stimuli. The University of Pennsylvania Smell Identification Test (UPSIT; Richard L. Doty, Shaman, Kimmelman, & Dann, 1984) is a clinical test of olfactory ability and uses 40 microencapsulated 'scratch and sniff' odorants within a standardized test of olfactory function. The creation of this test includes normative data for familiarity, pleasantness, intensity, and irritability, and has been used extensively in olfactory research (Nguyen, Ober, & Shenaut, 2012). However, the UPSIT is a test of olfactory dysfunction, where normal olfactory function would see naming of these highly familiar odours at, or near, ceiling. Employment of such a stimulus-set would provide limited variability in terms of familiarity and, potentially, encourage a memory strategy utilizing verbal labels. An alternative is to use odorants from the MONEX-40 (Freiherr et al., 2012), a test designed to detect differences in olfactory identification abilities in a normal population. However, the normative ratings from this study again focus only on familiarity, intensity, and pleasantness, and are limited to a relatively small set of 40 odorants.

Perhaps the closest attempt to a normative database for olfactory recognition tasks was reported by Sulmont et al. (2002). In this study, odours were rated in terms of familiarity, perceived complexity, and pleasantness by 24 French-speaking participants. Verbal identification was tested by selecting the name from a 68-item forced-choice list. These ratings were used to generate two familiar and two unfamiliar recognition sets of 18 odours. Interestingly, some perceptual overlap between dimensions was found with a significant positive correlation between pleasantness ratings and familiarity (R^2 = .53), a negative correlation between complexity and familiarity (R^2 = .65), and a positive correlation between notes (a different dimension of complexity) and familiarity (R^2 = .30).

Further to the primary aim of providing a database of olfactory normative data, the present study aims to advance the use of normative databases in olfactory memory research in two ways. First, whether subjective perceptual ratings of odours are suitable for use in a normative database is considered. Individual differences are undoubtedly present in the perception of odours (Kaeppler & Mueller, 2013), and are perhaps more influential than for perception of verbal or object stimuli because of strong top-down influences on odour perception (Wilson & Stevenson, 2006). If these individual differences exceed the differences obtained across the corpus of stimuli, it would suggest that tailoring odours to participants based on their ratings (Rabin, 1988) is a more effective method for stimulus control. Second, the relationships between the dimensions within this database are considered. As discussed in detail below, perceptual measures of olfactory stimuli are rarely independent, and the relationships between these dimensions should be considered when selecting stimuli for further tests.

2.1.1.1 Selection of perceptual dimensions for study

The present study involved the collection of normative data across a large set of commercially available odours (food and non-food odours are used since A N Gilbert & Greenberg, 1992 suggest that using food-related odours only may limit generalizability). A large number of measures were selected based upon past work with odours and different modality normative databases. Scales and questions were presented without accompanying interpretation guidance. That is, participants were free to interpret each measure as they wished. The justification for these measures is outlined below.

Verbalisability. The first dimension concerns the extent to which odours can be named. Typically, variations in odorant nameability have seen important effects on recognition (R. A. Frank et al., 2011; cf. Zucco, 2003), and dissociated neural activations for odours that can or cannot be named are suggested to reflect a dedicated mechanism for processing hard-to-name odours (Zelano et al., 2009). However, the name for an odour is an arbitrary construct which can include the source of the odour, a manufacturer name, or even a similar odour source it resembles. In addition, identification (and thus naming) of even familiar odours is often very poor (Lawless & Engen, 1977). As such, correct identification (the 'veridical label') is likely not important when considering the effect naming has on recognition, and its use for categorization may lead to an overestimation of the amount of 'un-nameable' odours. Rather, any odour that has an identifying verbal label attached to it should be considered as utilizing verbal codes (and could conceivably be represented as a verbal, rather than olfactory, code), whilst only very broad categories, such as a basic hedonic label, should be classed as non-verbalisable (Jönsson et al., 2011). In the present task participants are required to attach any verbal label to each odour, which is then scored according to the specificity of this label. However, a caveat to using the quality of labelling as a verbalisability measure is that consistency of labelling is not considered. That is, the naming of odours may only be important in memory experiments if the names attributed to the stimuli are consistently reproduced (Frank et al., 2011; Cornell Kärnekull et al., 2015). Despite this, a clear effect of this assessment of labelling quality has been observed on working memory performance (Jönsson et al., 2011) and thus appears to provide a reasonable measure of the role of verbal labelling in memory.

Describe-ability. Participants are also required to rate each odour's describe-ability (on a 7-point scale). Participants typically exhibit over-confidence in their ability to correctly name odours despite poor naming performance (Jönsson & Olsson, 2003). Discrepancies between participants' perceived and actual ability might reflect the difficulties in accessing the name of an odour; a feeling of knowing termed the 'tip of the nose' phenomenon (Lawless and Engen, 1977). However, the verbalisability score used in the present study is clearly a much more liberal criterion than odour naming ability. Since there is no 'wrong' verbal label, ability to label the odour is perhaps likely to reflect the participants' awareness of an odour's description (which would presumably include labels). Thus, with this method a strong relationship might be expected between perceived descriptive ability and actual ability to generate verbal labels.

Context availability. The third dimension is context availability. This measure is closely related to concreteness (Altarriba, Bauer, & Benvenuto, 1999), and refers to whether the odour can be easily associated to the context or circumstances with which the odour might appear. Whilst one might label this dimension imageability (i.e. the ability of the stimuli to evoke a mental image, John T. E. Richardson, 1975), such a label is avoided to prevent conflation with perceptual imageability (i.e. imagining the perceptual experience of an odour, see Stevenson et al., 2007).

The measure of context availability requires a 7-point rating of the ability to imagine the odour source. For example, the odour ‘lemon’ may evoke an image of a lemon, or the odour ‘chlorine’ may evoke an image of a swimming pool. For the latter, the odour (or in this case the context in which the odour is experienced) may be clearly imageable despite a poor ability to identify a source. It is possible, however, that this rating might again simply reflect the verbalisability of the odour, since an image is likely to result from the word that is associated with the odour.

Pleasantness. The hedonic rating of an odour features in many studies of odour perception and memory (Dalton, Maute, Oshida, Hikichi, & Izumi, 2008; Richard L. Doty, Shaman, Applebaum, et al., 1984; Nguyen et al., 2012; Sulmont et al., 2002). These studies show that pleasant/unpleasant odours result in activations in dissociated brain regions (Rolls, Kringelbach, & De Araujo, 2003), and are a particularly pertinent factor in odour perception by non-experts (Yoshida, 1964). Hedonic determination is considered a key function in olfaction and is even suggested to represent the primary method of discrimination between odours (Schiffman, 1974). Importantly for odour recognition tasks, less pleasant odours have produced better overall recognition (Nguyen et al., 2012), indicating an important role of the dimension in how odours are represented in memory. This finding also makes it important to match pleasantness of odours when inspecting the effects of other dimensions on recognition. In the present study, participants are required to rate pleasantness on a 7-point scale.

Intensity. The fifth dimension, intensity, is also measured on a 7-point scale. Although perceived intensity of an odour is related to the concentration of the odorant (Berglund, Berglund, Ekman, & Engen, 1971; Cain, 1969), it is also suggested to depend on experience-dependent factors (Ayabe-Kanamura et al., 1998; Distel et al., 1999). Specifically, the proposed degree of independence between intensity and pleasantness has varied from being entirely separate (Bensafi, 2002), to being related (Distel et al., 1999), or identical (Henion, 1971) constructs. Some studies have attempted to mitigate cross-condition differences in odour intensity by manipulating substance quantity (Stevenson et al., 2007) or via dilution (Sulmont et al., 2002). However, the odour intensity in the present experiment was allowed to vary between each odour, allowing investigation into its relationship with other factors across a broad range of intensities.

Irritability. The sixth dimension, and one potentially related to both intensity and pleasantness is the perceived irritability. An irritability measure is included in the normative data for odours in the UPSIT, and this measure would be expected to show a clear negative correlation with pleasantness as an additional reflection of a hedonic response. Irritability and pleasantness have shown differing effects on memory, where a recognition advantage for highly irritable odours is observed in older adults only (Larsson et al., 2009). Additionally, irritability has been used as an independently rated dimension when controlling high and low familiarity odour sets in memory tasks (Savic & Berglund, 2000). Whilst studies that do test irritability fail to clearly define this dimension, such a rating scale is likely interpreted as the physiological reaction to the odour. The findings by Larsson et al. (2009) indicate that a 7-point rating scale (very soothing/very irritating) will reveal a dimension that is independent of both pleasantness and intensity ratings.

Familiarity. The seventh measured dimension is familiarity. Odour familiarity is commonly a self-rated measure, though for verbal stimuli Brown and Watson (1987) suggest that subjective familiarity ratings are not a good substitute for objective frequency measures. This is because other factors such as frequency and age of acquisition ratings were found to contribute to judgments of familiarity (Brown and Watson, 1987). Despite this, such ratings of familiarity have been shown to be relatively stable when measured across different participants and time periods. For instance, ratings of familiarity from the UPSIT (Doty et al., 1984b) were utilized almost thirty years later in an odour memory study from Nguyen et al. (2012), and shown to correlate with new participant ratings (r = .46, p = .004). Similarly, Köster, Degel, & Piper (2002) compared familiarity scores provided for 12 odours with an earlier study (Degel et al., 2001) and found no significant differences in familiarity ratings.

Frequency. Familiarity is a complex construct which may be influenced by other dimensions. For example, word frequency is considered one of the most important variables in word processing (Brysbaert, Warriner, & Kuperman, 2014) and can be measured both objectively, via written or spoken appearances, and subjectively, via ratings of how often a particular word is experienced (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004). The eighth dimension included is therefore of odorant frequency. Whilst an objective frequency measure for odours might, theoretically, be possible, subjective self-ratings are a more practical method of assessment. Such a rating scale is demonstrated with verbal stimuli to be a valid, and at times better, predictor of recognition performance than corpus frequency (Balota et al., 2004). Previous work by Sulmont et al. (2002) suggests that frequency and familiarity may be closely related (R^2 > .85, p < .001). The present study will examine this through a 7-point rating scale.

Age of acquisition. A further construct that may influence familiarity (and the ninth measure in this study) is age of acquisition. Such a scale has not been studied previously for odours. It has, however, been shown to predict familiarity ratings and processing speed (Brown & Watson, 1987) for verbal stimuli. Age of acquisition for words is ideally mapped objectively by testing children on their naming ability, but has often been substituted for adult estimates of the age at which they first learnt the word. Morrison et al. (1997) suggest that these estimates can be a reliable and valid alternative measure if ratings (for example, because the sample are children) are unavailable. These age of acquisition ratings will allow a first examination of odour age of acquisition and explore the relationship with familiarity ratings. Participants will simply state the age at which they first experienced the odour. Instances where participants believe an odour to be novel will be coded as the current age of that participant.

Perceived complexity. The tenth and final dimension assessed in the present study is perceived complexity. Perceived complexity will be measured subjectively, since analysis of the chemical complexity of odours have shown no relationship to their perceived complexity (Jellinek & Köster, 1979). Subjective complexity ratings were shown to be reliable in a follow-up experiment, and as such are suggested to provide a meaningful measure in non-experts (Jellinek & Köster, 1983). One might expect that ratings of an odour's complexity would relate to the perceivable odours that combine to make it; however, Sulmont et al. (2002) suggest there may be separable dimensions of complexity ratings and the perceived odour notes in an odour. They propose that perceived complexity is related to familiarity of the item, with complexity ratings reflecting the extent the stimuli can be interpreted as a meaningful unit. That is, a familiar odour will be rated as more simple. This is supported by a clear negative correlation of complexity with familiarity ratings. Alternatively, Jellinek and Köster

(1983) have previously shown no relationship between complexity and familiarity, but used a measure of 'odour components' rather than a simple-complex rating scale. This question is presumably similar to the odour note question in Sulmont et al. (2002). It may be that an independent finding regarding 'odour notes' comes from the ambiguity of this question for naïve participants. As such, complexity ratings in the present study focus on a scale of rated simplicity/complexity, on a 7-point Likert scale.

2.1.1.2 Predictions

In utilizing a large number of odours in this normative study, the aim was to obtain a wide range of scores across the dimensions. Across these dimensions, some interrelation is expected. Previous work (Sulmont et al., 2002) reported positive correlations between pleasantness and familiarity and a negative correlation between complexity and familiarity. Intuitively, one might expect correlations between measures of verbalization and prior exposure (e.g. familiarity, frequency, and age of acquisition); with the necessity for labels developing if one regularly encounters the stimuli. It is also prosaic to predict a negative correlation between pleasantness and irritability. This is the first study to try and assess age of acquisition (i.e. first exposure) for odours. However, if age of acquisition effects emulate that of verbal stimuli (see Morrison et al., 1997), one might expect age of acquisition to correlate negatively with familiarity, frequency, and context availability (i.e. the earlier that one is first exposed to the odour, the higher the ratings of familiarity, frequency, and imageability). Intensity is also expected to relate to pleasantness ratings, either as an increase in intensity as odours are rated unpleasant (Sezille, Fournel, Rouby, Rinck, & Bensafi, 2014), or perhaps an increase in intensity as pleasantness deviates from neutral (hedonic strength, Distel et al., 1999).

2.1.2 Materials and methods

2.1.2.1 Participants

One-hundred and three non-smoker students (18 male and 85 female, mean age = 19.4, age range = 18-34) were recruited via Bournemouth University's online experiment management system, and participated for course credits. Participants who self-reported olfactory impairments (for example, symptoms of cold) were excluded, as were participants aged older than 40 years. Age-based exclusion was due to the proposition that olfactory identification abilities peak between the third and fifth decade (Richard L. Doty, Shaman, Applebaum, et al., 1984; see also Wood & Harkins, 1987 for age-related differences in the recognition of odours). Three female participants withdrew from the study after the first session, leaving usable data from one hundred participants. This study was carried out with approval from the Bournemouth University ethics panel. All participants gave written informed consent in accordance with the Declaration of Helsinki.

2.1.2.2 Design

A correlational design was used. The odours were grouped into 4 batches (A-D) of 50 odours (each containing 25 food and 25 non-food odours). Participants rated two of the four batches (that is, 100 odours) across two 60-minute sessions separated by a minimum of 24-hours. The presentation order of these batches was counterbalanced such that the testing orders A-B, B-A, C-D, and D-C were balanced across participants

2.1.2.3 Odorants

Two-hundred commercially available odorants (100 food-related and 100 non-food-related: see Appendix A for a complete list) were prepared by Dale Air Ltd. (www.daleair.com). These were stored within small test-tubes containing approximately

5ml of a liquid odorant soaked into a small piece of gauze. Due to contamination, odorant 17 (cabbage) was removed after 29 participant ratings. It remains included in the final database, but use of ratings for this odorant should be considered with caution.

2.1.2.4 Procedure

Testing was undertaken in a well-ventilated and quiet laboratory. Participants were tested in groups varying in size from 2-8. In the test phase, odours were presented on test-tube trays containing a block of five odours, with each odour arbitrarily numbered from one to two-hundred. Within each testing session participants received 10 blocks of 5-odours, meaning participants smelled 50 odours in each of the two sessions. The composition of each 5-odour block was selected at random from the odour set within each batch. Participants were instructed to evaluate those odours in any order.

Evaluation required participants to open the test tube lid and smell the odour (birhinally) for approximately 3 seconds in order to answer each question. Between odours, participants took a break of approximately 20 seconds, and between odour blocks a break of 1 minute was implemented where participants would take a drink of water. Responses were recorded within a booklet wherein each odour was assessed across the 10 dimensions. Ratings were measured on a 7-point Likert scale, labelled at each end, and at the neutral centre point. Each dimension (identified from the literature discussed above) was presented in the same order for each odour and participant. Participants were asked: ‘How *familiar* is this odour (not at all familiar/very familiar)’, ‘how *intense* is this odour (very weak/very intense)’, ‘how *pleasant* is this odour (very unpleasant/very pleasant)’, ‘how *complex* is this odour (very simple/complex)’, ‘how *irritating* is this odour (very soothing/very irritating)’, ‘how *frequently* is this odour experienced (not at all frequently/very frequently)’, ‘how easy is it to *describe* this odour (very difficult/very easy)’, and ‘how easy is it to *imagine* where you’d experience this odour (very difficult/very easy)'. In addition, two questions were open-ended. The first required a numerical age of acquisition response to 'at what age did you first experience this odour?', and the second a verbal written response to 'can you attach any labels to this odour?' Participants were instructed to rate independently and in silence, and, if uncertain, participants were asked to guess.

2.1.3 Results and discussion

2.1.3.1 Scoring protocol

The first eight questions were coded on scales of 1-7 (familiarity, intensity, pleasantness, complexity, irritability, frequency, perceived describe-ability, and context availability).

In reporting age of acquisition, participants were encouraged to estimate the age at which an odour was first encountered, and provide a single age. When participants reported an age range as their answer, the median value of that range was recorded. A small number of participants provided a qualitative (rather than quantitative) age of acquisition response (for example, "childhood"). In this instance the age of acquisition score was not used.

The scoring of odour labels (verbalization) followed a modified version of the method described by Jönsson et al. (2011). These labels were coded on a 4-point scale (0-3). No response or a very basic affective judgment received a score of 0. Broad categorizations or generic labels (for example; cleaner, food, sweet) received a score of 1. More specific categorizations referring to specific groups (floral, perfume, sweets) received a score of 2, and any specific noun label received a score of 3. Scoring was performed independently by two researchers, with the median score taken as the final verbalisability value. Weighted Cohen's κ determined a good (Altman, 1991) level of agreement between raters, $\kappa_w = .61$ (95% CI, .59 to .62), $p < .0005$.

Responses were averaged across participants to give a normative score in every odorant for each dimension. The full normative ratings for the 200 odours can be found in Appendix A.

2.1.3.2 Normative data reliability

In order for the normative data to be transferable to other samples in future studies, it is important to demonstrate that the variance in the ratings is attributable to the odours rather than individual differences in perception of the odours. Should the variance across participants match or exceed the variance between odours, it would suggest that tailoring odours according to individual participant ratings would be more suitable (Rabin, 1988).

In order to test this proposition each dimension was looked at individually, using an analytical method described by Uebersax (2015). The agreement of scores for each odour across participants (individual differences) was examined as a measure of variability. That is, for each dimension, an individual's rating of each odour was correlated with the average rating for that odour (a measure of 'consistency across participants'). The higher the correlational coefficient, the greater the agreement between raters. Conversely, the lower the correlational coefficient, the greater the individual differences between raters. To assess the discriminability between odours, each individual's rating of an odour was correlated with their average rating across all odours for that dimension. A high correlation coefficient (a measure of 'consistency across odours') indicates little variation in the scores given for that dimension by each participant across odours. That is, participants respond similarly for that dimension across odours, indicating that the dimension is weak in discriminating between the odours. For the normative data in each dimension to be considered suitable the effect size for odour score agreement should significantly exceed that of rater score agreement. That is, ratings for an odour on each dimension should have a stronger relationship with the mean rating for that odour compared to the relationship to the mean rating across odours. A series of t-tests were conducted to test this proposition, comparing the strength of effect size for the odour (consistency across participants) and the level of discriminability (consistency across odours) for each of the dimensions, and is shown in Table 1. As can be seen from the table, the effect sizes for these relationships differ across dimensions, so require some further consideration.

For ratings of familiarity, pleasantness, irritability, and intensity, the association of participants' responses to the mean response for an odour was significantly greater than the association of responses to the mean response for each participant. That is, responses for a particular odour were more closely associated to the normative score for that odour than they were to each participant's average response on that dimension. This suggests that those four dimensions are capable of discriminating between odours above any general response bias/strategy applied to that dimension. For complexity and age of acquisition, participants' ratings were more strongly related with the average rating for that dimension. This suggests a lack of sensitivity for complexity and age of acquisition. This finding may be due to several reasons. Participants may have shown little variability in how they respond for each odour, resulting in each response showing a strong relationship with the mean. For example, if they are unable to conceptualize 'complexity' and 'age of acquisition' they may adopt a default response for the question resulting in limited variability. Alternatively, a low association of ratings to each odour mean indicates a large effect of individual differences. Indeed, Table 1 shows that age of acquisition and complexity exhibited the lowest consistency across participants, indicating greater individual differences. These individual differences could occur through genuine variation in the ages at which an odour is first experienced or in the perceived complexity of odours, though may also arise from participant difficulties in interpreting and applying the particular question to the stimuli. Furthermore, ratings for frequency, context availability, and describeability, in addition to the labelling scores, showed no significant differences between the consistency across participants and consistency across odours. Consequently, these dimensions may exhibit reduced discriminatory power within a normative database.

Table 1

Mean (SD) r coefficients of rater agreement with each odour's mean score, and rating agreement with each rater's mean score.

	Dimension									
	Fam.	Int.	Pleas.	Comp.	Irr.	Freq.	Desc.	CA	AoA	Verb
Consistency across participants	.484 (.09)	.484 (.10)	.611 (.10)	.263 (.14)	.563 (.11)	.422 (.11)	.421 (.11)	.427 (.10)	.373 (.11)	.411 (.14)
Consistency across odours	.422 (.12)	.406 (.12)	.312 (.14)	.408 (.12)	.408 (.13)	.434 (.12)	.417 (.12)	.433 (.12)	.512 (.13)	.435 (.13)
***t* value**	4.69*	5.56*	21.69*	-9.03*	10.37*	-0.16	0.26	-0.42	-9.32*	-1.48

* Comparisons significant to $p < .001$.

2.1.3.3 Relationships between dimensions

The linear correlation coefficient (*r*) was calculated for each dimension pairing, and displayed as a correlation matrix in Table 2. Almost all correlations were significant, with the exception of the intensity dimension with familiarity, frequency, describe-ability, context availability, and age-of-acquisition dimensions.

Some of the dimensional correlations warrant additional comment. As noted in the introduction to this chapter, this is the first study to attempt to assess the effect of age of acquisition in olfactory processing. Consistent, with the verbal domain (Morrison et al., 1997), age of acquisition displays strong negative correlations with familiarity, frequency, and context availability. As expected, a strong negative correlation between age of acquisition and labelling was also reported, suggesting early exposure provides increased opportunities in which to develop a label for that odour. Indeed, inter-correlation was observed for several dimensions relating to knowledge and previous experience with the odorant. The strong relationship is present between these dimensions despite evidence that individual differences may exceed the variation observed across odours. Consequently, it is possible that these ratings may still have utility in a normative database, aiding researchers in odour selection before further tailoring of stimuli according to participant data.

Of particular interest are the four dimensions identified as particularly suitable for use in a normative database; familiarity, pleasantness, irritability, and intensity. First, the strong negative correlation (r = -.98) observed between irritability and pleasantness suggests collinearity, so further discussion focuses on only pleasantness scores. A predicted positive correlation between familiarity and pleasantness (Sulmont et al., 2002) was observed, and supports a classical mere-exposure effect (Zajonc, 1968). Also predicted was a linear negative relationship between intensity and pleasantness (Sezille et al., 2014), or a non-linear relationship where intensity increases with both pleasantness and unpleasantness (Distel et al., 1999). In the present data, though a linear model was significant, $F(1, 198) = 75.23$, $p < .001$, $R^2 = .28$, a quadratic curve better fit the data, $F(2, 197) = 88.16$, $p < .001$, $R^2 = .47$ (Figure 1A). When pleasantness data were recoded as a measure of hedonic strength (with neutral responses scored as 0, increasing to 3 as they deviate above or below neutral), a linear model was accepted as the best fit, $F(1, 198) = 181.20$, $p < .001$, $R^2 = .48$ (Figure 1B). That is, intensity ratings are linearly related to the strength of a hedonic response. A strong relationship between hedonic strength and intensity supports ideas that the two may reflect similar dimensions of odour judgment (Henion, 1971).

However, a non-significant relationship between intensity and familiarity (Figure 1C) is an interesting result that is not consistent with the findings in Distel et al. (1999). They suggested that not only might an increased familiarity with a stimulus affect judgments of pleasantness (a relationship seen in these data, Figure 1D), but also that intense odours may be more easily recognized and thus more likely to be judged as familiar. The observed pattern of relationships between familiarity, pleasantness, and intensity instead suggest that familiarity and intensity contribute independently to pleasantness scores.

Table 2

Correlation matrix (r) of averaged scores across participants for each odour.

		Q1.	Q2.	Q3.	Q4.	Q5.	Q6.	Q7.	Q8.	Q9.	Q10.
Q1.	Familiarity	—									
Q2.	Intensity	.05	—								
Q3.	Pleasantness	.73*	-.53*	—							
Q4.	Complexity	-.40*	.64*	-.63*	—						
Q5.	Irritability	-.68*	.61*	-.98*	.66*	—					
Q6.	Frequency	.92*	-.08	.77*	-.50*	-.73*	—				
Q7.	Describeability	.94*	.09	.67*	-.42*	-.62*	.92*	—			
Q8.	Context Availability	.95*	.09	.66*	-.40*	-.61*	.93*	.97*	—		
Q9.	Age of Acquisition	-.91*	.03	-.72*	.45*	.69*	-.88*	-.89*	-.90*	—	
Q10.	Verbalisability Score	.88*	.15*	.54*	-.28*	-.50*	.82*	.88*	.90*	-.86*	—

* Significant correlations at the 0.05 level.

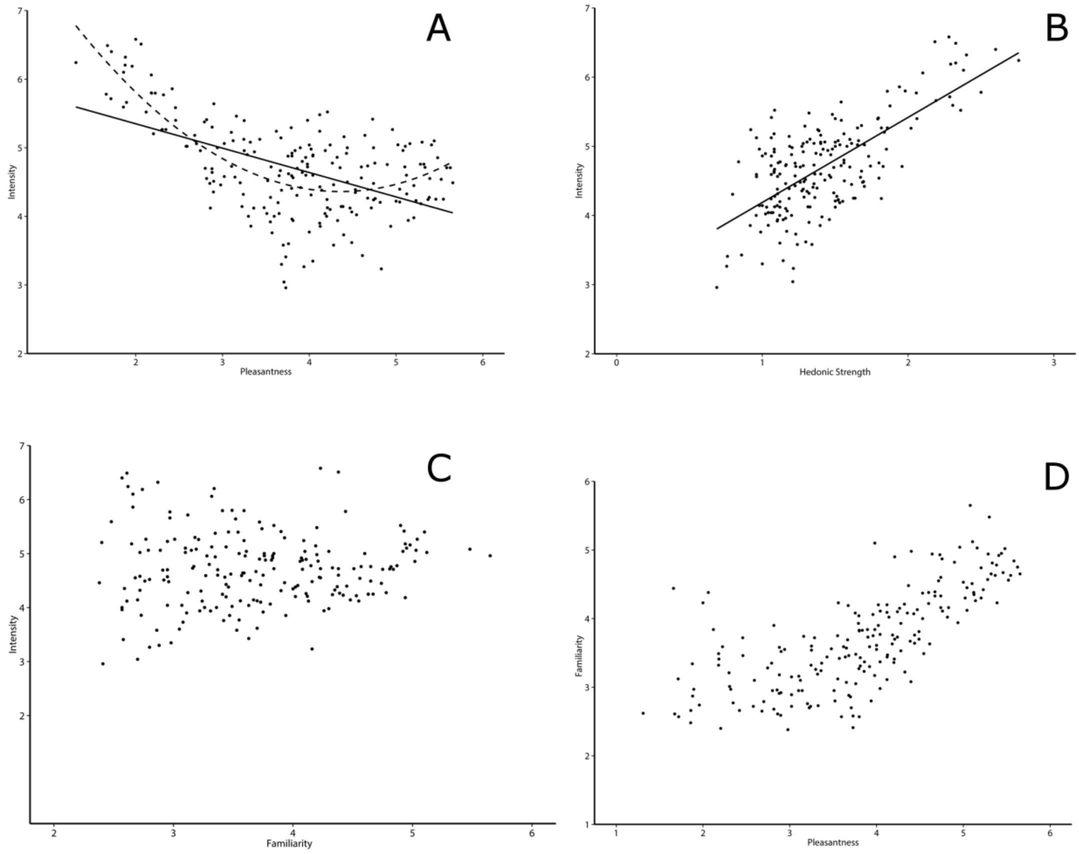

Figure 1. Relationship of (A) pleasantness with intensity, (B) hedonic strength with intensity, (C) intensity with familiarity, and (D) familiarity with pleasantness.

2.1.4 Discussion

The present study provides a large-scale normative dataset, containing ratings from 10 dimensions for 200 commercially available odours (see Appendix A). To date, this is the largest such study examining psychological dimensions for olfactory stimuli. These ratings are available in Appendix A and provide the necessary information for researchers to control dimensions in subsequent studies (indeed, these data are used in this thesis for controlling olfactory stimuli). Additionally, these normative data are the first to assess the effects of age of acquisition on olfactory processing. Whilst similar relationships with age of acquisition are shown with words (i.e. there is a strong negative correlation between age of acquisition and familiarity, Morrison et al., 1997), age of acquisition was shown to be strongly influenced by individual differences and the dimension did not adequately differentiate between odours. As a result, the increased unsystematic variance in the age of acquisition norm values means they should be used with caution. Moreover, it is interesting to note that age of acquisition also exhibits strong negative correlations with frequency, describeability, context availability, and labelling score. Consequently, attempts to use the norms to isolate any effects of age of acquisition may be confounded by these inter-relations.

Normative data are suggested to provide two important benefits. The first benefit concerns experimental control. Since memory for odours has been shown to be affected by factors such as familiarity (e.g. Yeshurun et al., 2008) and pleasantness (Nguyen et al., 2012) it is argued that this may be of utility to control for such issues, analogously to that done with verbal memory. For example, if one were comparing memory for odours across two conditions (e.g. under conditions of quiet and concurrent articulation), matching the odours using these dimensions would eliminate a possible confound in that comparison. More specifically, studies examining serial position effects for odours report both differences across studies and potential qualitative differences with the functions reported for other stimulus types (Reed, 2000; Miles and Hodder, 2005; Johnson and Miles, 2007; 2009; Johnson et al., 2013). It is possible that these differences may be the effects of irregularities in the selection of stimuli; indeed, Hay et al. (2007) suggest that the psychological distinctiveness of stimuli can affect the shape of serial position curves. This study provides a database from which researchers can systematically examine whether such serial position effect differences can be explained by characteristics of the odours. However, it should be noted that the data highlights some caveats in the selection of these dimensions since only the normative ratings for familiarity, pleasantness, irritability, and intensity exhibit convincing discriminatory power. If researchers intend to investigate the effects of the remaining dimensions, it may be advisable to follow the approach undertaken by Rabin (1988), i.e. tailoring odours to individual ratings.

The second benefit of this normative dataset is that it provides a framework from which other researchers can examine the effects of psychological dimensions on olfactory memory. Researchers can use these data to explore whether dimensions that affect verbal memory similarly affect olfactory memory (as these odours are commercially available). One might expect that manipulating the familiarity of the stimulus set using these data would be of most interest in order to compare perceptual memory and the potential facilitative effects of verbal-perceptual dual-coding (Yeshurun et al., 2008).

That intensity was allowed to vary arguably reduces the usefulness of the normative data to the specific stimulus set used. It is possible that the relationship of intensity with pleasantness, and to some extent with irritability and complexity, may confound the scores obtained for these dimensions. This is not considered a particular limitation, as the aim of the present study was to provide these data for a stimulus set that is readily-available and which does not require researchers to manually match odorant intensities to n-butanol. Selecting odorants for future research from the database can include matching odorants on intensity, whilst still allowing dimensions of interest to be manipulated. Furthermore, although several odours are artificially produced to reflect non-tangible objects (e.g. ‘sports locker room’), many of the odours are labelled from real-life objects. There is therefore opportunity for future research to expand the utility of these data by comparing other odour sources with the normative scores presented here.

The normative data may be, to some extent, limited by the sample. The majority of participants were female (85%) and, in general, when sex differences are found in respect to olfaction females exhibit superiority (see Richard L. Doty & Cameron, 2009 for review; although this trend can be complicated by menstrual cycle; e.g. Doty, Snyder, Huggins, & Lowry, 1981; Purdon, Klein, & Flor-Henry, 2001). Of particular relevance to these normative data is the finding that females exhibit superior identification of odours (Larsson et al., 2004). Indeed, Öberg, Larsson and Bäckman (2002) have shown that when naming ability is controlled, sex differences are removed (see also Larsson, Lövdén, & Nilsson, 2003). That females are superior at naming odours may result in an inflation of the verbalisability score for the odours. Similarly, the use of university students in the sample may also have led to an overstatement of the name-ability of the odours. This is because educational level has been found to be a reliable predictor of odour identification (Moberg et al., 2014). Whilst this sample may have resulted, quantitatively, in a general inflation of ratings (particularly with respect to odour naming), there is no a priori reason to suggest that perception of these odours may have changed qualitatively with more males or a less educated sample. Consequently, it is argued that the relative differences between the odours remains and the data retains its utility in differentiating odours. Notwithstanding, it is possible that these norms, particularly for food-based odours, may be limited cross-culturally. Gilbert and Greenberg (1992) question the universality of food-related odours since “what smells like food to persons of one culture may not smell edible to those of another” (p.327). Different experiences with odours across cultures, both qualitatively and quantitatively, may fundamentally change conceptualization of those items. As a result, these norms may not translate to other cultures; although this is an empirically testable question that warrants further examination.

One might argue, however, that restricting the sample to a British-born student population functions to limit individual differences in the ratings of the odours, e.g. less culture-based variance in the preference for food-based odours (Kaeppler & Mueller, 2013). Notwithstanding this limitation in sample variance, some dimensions are identified that are less suitable for use in normative databases due to high levels of individual difference and/or a lack of sensitivity in discriminating between odours. For these dimensions a participant's average response across odours is more predictive of the rating than the average rating for that odour. This suggests dimension insensitivity. For these dimensions there was either a high level of variability between participants, or participants were conservative in the spread of scores they gave each odour. Interestingly, it is the dimensions that are most commonly considered in olfactory research that demonstrated most suitability for use in normative databases (those of familiarity, intensity, and pleasantness/irritability). However, the scales that did not meet the criterion of agreement should not be discounted. For example, the verbalisability scale was designed based on previous n-back research (Jönsson et al., 2011), and has shown working memory differences for odours selected based on this score. Further, correlations demonstrated between normative scores across dimensions, particularly those that have previously demonstrated relationships, support the validity of these scores. Therefore, rather than claiming that some dimensions lack utility, the data suggest that for some dimensions, individual differences/response biases may create more unsystematic variance in the normative values.

In summary, the normative data presented here may be utilized in future research to control odours for differences in olfactory perception. The dimensions should, however, be used with consideration of individual differences, particularly if testing a dissimilar population to that tested here. The ratings presented here do not offer a replacement for tailoring odours to participants (Rabin, 1988), but should be used where prior exposure of odours to participants is not desirable, or used to guide selection of odorants which can be later supplemented by post-hoc rating and categorization.

Chapter 3: Investigating olfactory working memory using the n-back task

3 Chapter Summary

A series of studies which replicate and extend the n-back experiment in Jönsson et al. (2011) are described. The aim was to investigate the relative contributions of verbal and perceptual coding to olfactory working memory (as measured by n-back performance) and discriminability. The facilitative effect of verbal labelling on odour working memory was examined across two testing sequences, revealing a working memory advantage for verbalisable odours in only the second sequence (Experiment 1). This was attributed to verbal learning that improved discriminability, accentuated use of a verbal rehearsal strategy, or both. The use of a rehearsal strategy was explored in Experiment 2 using a dual-tasking procedure, where it was attempted to limit the use of verbal or visual codes during n-back maintenance. There was evidence against (using Bayes Factors) a general working memory deficit, or attenuation of the high verbalisability advantage, during concurrent articulation. This contradicted the proposition that verbal rehearsal of odour labels underpinned the n-back advantage for these odours. Experiment 3 applied the remember-know paradigm to examine the role of familiarity (an automatic strength signal) and recollection (controlled retrieval of contextual information) in olfactory n-back performance. This revealed a quantitative improvement in recollection for high verbalisability odours, and no difference in responses using familiarity-based processes. This finding suggests that the memory advantage for highly verbalisable odours is related to more recollection of the odorants. Finally, Experiment 4 sought to examine the role of familiarity on odour discriminability, by experimentally inducing perceptual familiarity to assess the subsequent effect on n-back performance.

An unexpected pattern of results was observed where familiarisation decreased general n-back performance; this is discussed in relation to strategy adjustments due to conflict between item familiarity and serial-position recollection.

Investigating olfactory working memory using the n-back task

3.0 Chapter Introduction

The traditional conceptualisation of working memory defines the ability to temporarily store, rehearse, and manipulate information (Baddeley & Hitch, 1974). Whilst a number of studies have sought to examine aspects of olfactory short-term memory (e.g. Andrade & Donaldson, 2007; Jehl, Royet, & Holley, 1997; A. J. Johnson, Cauchi, & Miles, 2013; A. J. Johnson & Miles, 2007, 2009; Miles & Hodder, 2005; Miles & Jenkins, 2000; Walk & Johns, 1984; White, Hornung, Kurtz, Treisman, & Sheehe, 1998; Zelano, Montag, Khan, & Sobel, 2009), there exists a stark paucity of work examining olfactory working memory (OWM) in line with this definition. Indeed, as discussed in Chapter 1, it is unclear to what extent memory for odours includes a conscious representation (Stevenson, 2009; Zucco, 2003) that may be manipulated or refreshed in working memory (Raye et al., 2007).

Both Dade et al. (2001) and Jönsson et al. (2011) employed a 2-back task, requiring maintenance and manipulation of a presented sequence of odours, in order to examine OWM. In their study, Dade et al. (2001) compared 2-back performance for faces and odours and reported similar performance levels for the two stimulus types (≈90%). In addition, similar activations in the dorsolateral and ventrolateral frontal cortex for both faces and odours suggested working memory operations that were independent of stimulus modality. However, as noted by Jönsson et al. (2011), the selected odours used by Dade et al. were highly familiar and, as a consequence, OWM performance may have been supported via verbal recoding and rehearsal of such labels. Following recoding, working memory performance may therefore have reflected verbal, rather than olfactory, representations. Indeed, there is evidence to suggest that the nameability of an odour affects memory performance (as well as the affective experience, De

Araujo, Rolls, Velazco, Margot, & Cayeux, 2005). This is shown quantitatively (Jehl et al., 1997; Lyman & McDaniel, 1986, 1990; Valentin, Dacremont, & Cayeux, 2011; Yeshurun et al., 2008) but also qualitatively in respect to patterns of neural activity (Zelano et al., 2009) and susceptibility to proactive interference (see Chapter 5).

Jönsson et al. (2011) addressed the issue of verbalisability in OWM by comparing 2-back performance for odours that had been categorised as high or low verbalisability. Whilst they observed superior n-back performance for the high verbalisable odours, performance for the low verbalisable odours remained above chance. This supports the notion that the perceptual code for an odour can be retained and updated within working memory (see also White et al., 1998 for evidence of olfactory perceptual representations). However, individual item analysis of n-back performance (*A'*) was strongly predicted (R^2 = 0.95) by the verbalisation score for each odorant, even in the hard-to-verbalise group. Consequently, there is support for a strong influence of verbal codes on olfactory n-back performance.

Notwithstanding above chance perceptual memory, Jönsson et al. demonstrate that odour verbalisation can improve task performance through increased discriminability, though the variance across odours was not fully explained by this discriminability advantage (see also Mingo & Stevenson, 2007; Stevenson, 2012, for effects of familiarity on discriminability). Indeed, whilst increased discriminability may provide one explanation as to why performance is superior for verbalisable odours, other explanations concerns the utilisation of perceptual-verbal dual-coding (Paivio, 1990; Stevenson & Wilson, 2007; Yeshurun et al., 2008; See Chapter 1 for a detailed discussion of the relationship between verbal and perceptual processing), or the access of a perceptual representation to consciousness facilitated by semantic information (Tomiczek & Stevenson, 2009).

3.0.1 The n-back procedure

The n-back task is popular in cognitive neuroscience due to the simple way in which task difficulty can be manipulated without changing the presentation style of stimuli or nature of participant responses (Dong, Reder, Yao, Liu, & Chen, 2015; Jaeggi, Buschkuehl, Perrig, & Meier, 2010). Participants are presented with a continuous sequence of stimuli, and a decision must be made whether the currently presented item matches the item *n* trials previously on a predetermined criterion. This criterion is typically the matching of item identity between probe and n-back item, but can also require matching the spatial location of items regardless of identity (Owen, McMillan, Laird, & Bullmore, 2005). A *target*, requiring a positive response, is a trial that matches that *n*th item back on this determined criterion. As *n* increases, the proposed load on working memory systems also increases, evidenced by an increase in reaction time and decrease in accuracy (Jonides et al., 1997; B McElree, 2001). Importantly, the nature of the n-back task is described as a *maintenance plus manipulation* task when $n > 1$ (Ragland et al., 2002), due to the need to update stored information as the trials progress.

Non-target items are called *lures*[1], and are typically taken from the same pool of stimuli as target items. That is, the stimulus used as a lure item may be used later in the experiment as a target, or vice versa. Importantly, the task allows an assessment of

[1] It should be noted that in this thesis 'lure' refers to all non-targets, rather than just non-targets that are close to a potential target position (as used by Kane et al., 2007; Schmiedek, Li, & Lindenberger, 2009). When discussing these alternative lure items, the term 'recent-lures' is used.

olfactory working memory as it does not necessarily require explicit identification or naming of the stimuli (Jönsson et al., 2011). In contrast, a complex span procedure utilises a recall-based method for assessing working memory capacity (Daneman & Carpenter, 1980; Turner & Engle, 1989). This category of task is a commonly used measure of working memory capacity that involves presentation of a to-be-remembered sequence of items with a secondary distractor task completed during the inter-stimulus-interval. However, the recall aspect means application to olfactory stimuli would simply test memory for odour labels (e.g. see Annett & Lorimer, 1995, and the procedure outlined by Miles & Jenkins, 2000). Consequently, the n-back procedure is suitable for measuring a participant's ability to manipulate and store olfactory information in working memory, provided appropriate steps are taken to assess the use of executive control.

3.0.2 What does the n-back task measure?

The n-back procedure has received some criticism over its validity as a working memory measure. Though the n-back task has face validity as a working memory task, it has shown little correlation with complex span (Jaeggi, Buschkuehl, et al., 2010; Redick & Lindsey, 2013; Simmons, 2000), a task often considered the gold-standard for measuring working memory capacity (Shelton, Elliott, Matthews, Hill, & Gouvier, 2010). Complex span tasks are a commonly used measure of working memory ability due to their strong predictive ability for tests of higher-order cognition, such as fluid intelligence measured through reasoning tasks (Barrouillet & Lecas, 1999; Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Conway et al., 2003; Kyllonen & Christal, 1990; Nash Unsworth & Engle, 2007). This lack of concurrent validity is therefore problematic for models that suggest both n-back and complex span measure the same working memory construct (see Kane, Conway, Miura, & Colflesh, 2007).

Paradoxically however, despite the apparent disparity between complex span and n-back measures, a relationship between n-back performance and fluid intelligence has also been observed (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Kane et al., 2007; Schmiedek, Lövdén, & Lindenberger, 2014). Whether this predictive ability of both complex span and n-back performance on higher-order cognition supports a shared mechanism across the tasks, or independent contributions to intelligence, is somewhat equivocal.

Kane et al. (2007) found both tasks predicted independent variance in measures of fluid intelligence, and a meta-analysis by Redick and Lindsay (2013) supports a view that the two tasks cannot be used interchangeably as working memory measures. They suggest that this discrepancy instead supports a multi-faceted working memory system that includes non-unitary executive functions such as shifting, updating, and inhibition (Miyake et al., 2000; Oberauer, 2009). Indeed, a key difference between complex span and n-back tasks, and a possible reason for a weak relationship between the two, is the reliance on retrieval through recall in the former and through recognition processes in the latter (Harbison, Atkins, & Dougherty, 2011; Jaeggi, Buschkuehl, et al., 2010; Oberauer, 2005; Redick & Lindsey, 2013; Shelton et al., 2010). Specifically, the recognition process required for the n-back task is influenced by both familiarity and recollection processes (Jaeggi, Buschkuehl, et al., 2010; Kane et al., 2007). A familiarity signal (for example, through elevated activation in LTM, Oberauer, 2009) is found in recently-presented items, which on its own may be sufficient for accepting target items (Harbison et al., 2011; Kane et al., 2007). However, it is the inclusion of 'recent' lures (i.e. lures that have previously appeared in positions n-1, n+1, or n+2) that produces a familiarity signal comparable to target items, and means this signal alone is not sufficient to distinguish the targets from non-targets. Participants are therefore required to engage in a control process over a familiarity signal to determine the correct position of the stimulus (Harbison et al., 2011; Szmalec, Verbruggen, Vandierendonck, & Kemps, 2011).

This process of *control over familiarity* may obscure the relationship between the n-back task and recall-based working memory measures (Kane et al., 2007). This has been supported by implementing a recall modification to the n-back (Shelton, Metzger, & Elliott, 2007; Wilhelm et al., 2013), which has shown a stronger relationship with complex span performance (e.g. r = .32 with operation span and r = .41 with listening span, Shelton et al., 2010). In this procedure, participants are presented with multiple sequentially-presented lists of items that vary in length, and at the end of each list are instructed to report the 1, 2, or 3-back item. Recall performance of 2 and 3-back items is then used to index working memory ability. Consequently, the task tests maintenance and manipulation without requiring a process of matching a probe to the n-back position.

Alternatively, the discrepancy between these working memory measures may be due to the design of the n-back procedures. Harbison et al. (2011) suggest n-back and complex span tasks show a weak relationship because in most demonstrations of the n-back task, approximately 50% of trial items appear as ‘non-recent’ lures. This means that a large proportion of an n-back score, based on an index of hits and false alarms, is made up of lure items where the participant need not attempt recollection. However, it should be noted that Kane et al. (2007) compared n-back scores based on only recent-lure performance with operation span, and found only weak correlations between the two. That is, for n-back lure items where control over familiarity-based responding was challenged, a weak relationship with complex span tasks remained. To be clear, though the reliance on item familiarity for these lure types may complicate the assessment of the relationship between the n-back and complex span, the inclusion of these trials does not appear to be the primary reason for any disparity.

Though there are clearly issues with treating n-back performance as working memory capacity analogous to that measured in complex span tasks, it is unclear whether they reflect unrelated constructs in working memory. Indeed, the differences across measures might simply reflect paradigm-specific variance which systematically reduces the observed relationships between the two tasks (Schmiedek, Hildebrandt, et al., 2009). This mismatch of memory-test methods, in addition to content-specific variance and measurement error, is proposed to be responsible for the low correlations between n-back and other tasks. The underlying working memory constructs may be better examined by assessing the relationship between latent variables based on multiple versions of both n-back and complex span (Schmiedek et al., 2014). This method has consequently revealed near-perfect relationships between the two category of tasks (Schmiedek, Hildebrandt, et al., 2009; Schmiedek et al., 2014; Wilhelm et al., 2013). Whilst correlations between individual tasks did vary considerably, latent variables of complex span (from measures of reading span, counting span, and rotation span) and the n-back task (using numerical and spatial n-back) correlated substantially (e.g. r = .69, Schmiedek et al., 2014). Furthermore, these variables loaded highly onto a working memory factor, which in turn was predictive of reasoning ability. The authors suggest a better measure of working memory can be achieved by using multiple heterogeneous tasks such as the complex span and n-back to produce a latent working memory factor.

The shared explained variance from complex span and n-back tasks indicate a crucial component of working memory capacity utilised in both procedures, and in measures of reasoning ability for which both tasks have predictive utility. Controlled attention, proposed as a domain-general process for maintaining and manipulating working memory items (Cowan, 1999; Engle et al., 1999), may be the source of this shared variance. Transfer effects from training in the n-back procedure to performance in reasoning tests is proposed to occur because attentional control is essential for both tasks (Jaeggi, Studer-Luethi, et al., 2010; Jaeggi et al., 2008). Similarly, attentional control is also considered essential for complex span tasks, as information about the stimulus must be accessible whilst attention shifts to the processing task (Engle & Kane, 2004; Kane et al., 2004). Alternatively, the common source of variance between complex span and n-back tasks may be the ability to create, maintain, and update bindings (Oberauer, 2005, 2009; Wilhelm et al., 2013). That is, the n-back task requires constant creation and updating of bindings between an item and its context, whilst complex span tasks have similar binding requirements where the item must be bound to its serial position (Wilhelm et al., 2013). In Wilhelm et al., a binding factor accounted for 100% of an updating factor's variance that included the n-back task, and 90% of complex span variance.

In summary, when variance from retrieval differences in the n-back and complex span task is accounted for, there is a strong relationship between the two tasks that suggests they share a common working memory function (Schmiedek et al., 2014; Wilhelm et al., 2013). Consequently, the n-back task appears to be a valid measure of working memory, though the influence of familiarity-based recognition on n-back performance should be considered when designing this working memory measure (e.g. Kane et al., 2007).

3.0.3 Strategy adoption in the n-back task

The n-back procedure is a complex task that may require (1) maintenance of the previous *n* items in memory, (2) updating of new items for active maintenance, (3) creation of bindings between each stimulus and its temporal context, and (4) resolving interference from non-relevant trials (Chatham et al., 2011; Cohen et al., 1997;

Oberauer, 2005). However, strategy adoption in the task can vary between participants or task demands (Botvinick, Braver, Barch, Carter, & Cohen, 2001), and the strategy used can mediate the working memory resources engaged (Juvina & Taatgen, 2007). The *Dual Mechanism of Control* framework (Braver, 2012) describes a proactive control process that involves activating the target n-back item in advance, ready for comparison to the anticipated stimulus. That is, for each trial the 'correct' n-back item is dynamically prepared prior to presentation of the item to which a response must be made. Specifically, participants activate the n^{th} back item in memory prior to the trial to decide if the forthcoming item is a 'hit'. In comparison, reactive control initiates attentive mechanisms upon presentation of the stimulus, after which competing responses are activated and the correct response must be selected. Specifically, following presentation of the item, participants attempt to retrieve the n^{th} back item in memory to establish if a correct match exists.

Several models have been produced that attempt to explain the precise processes involved in effective n-back task performance (e.g. Chatham et al., 2011; Gosmann & Eliasmith, 2015; Juvina & Taatgen, 2007; Szmalec et al., 2011). A common component in such models is an active rehearsal strategy, where a rehearsal window of size *n* is maintained (Harbison et al., 2011; Juvina & Taatgen, 2007; Szmalec et al., 2011). These rehearsal strategies involve *updating,* where new items are added to the rehearsal window and the now-irrelevant items removed (see also Wilhelm et al., 2013). However, rehearsal strategies may themselves differ in whether proactive or reactive control processes are employed (Ralph, 2014).

A proactive, *static rehearsal* process, is proposed in Chatham et al. (2011) where the *n* serial positions are held in memory separately to the *n* memory items. In this model, each item is allocated to a serial position, but only the task-relevant serial position is under the focus of attention when a new item is presented. This new item is then compared to the item in that serial position, a response is made, and the new item then replaces the previous memory item in that same position. Importantly, the attentional focus then allocates the next serial position as task-relevant, and the same matching process is completed. Such a method is efficient because it does not require updating of the position of every item, and does not require any updating should the new item match the old (a target) (Ralph, 2014).

In comparison, Juvina and Taatgen (2007) describe a *rolling rehearsal* strategy where items are rehearsed in the phonological loop to increase their activation strengths. At each trial, new items are appended to the list whilst the first item is removed. That is, when an item falls out of the maintenance window this item becomes irrelevant, and this highly activated item must be removed and inhibited to prevent its reappearance in the rehearsal list. However, interference may still arise from this removed item due to the limited capacity of a suppression mechanism (Harbison et al., 2011; Jonides & Nee, 2006; Juvina & Taatgen, 2007; Szmalec et al., 2011). This strategy reflects a reactive control process, where all the items in the rehearsal window are retrieved in response to a new stimulus, and the strongest activation in this list is then selected and compared to the presented item (Ralph, 2014).

Though the above strategies may differ in their use of proactive or reactive control, they are similar in their requirement for active maintenance of the stimulus. Importantly, it should also be noted that such maintenance is not necessarily specific to rehearsal within the phonological loop, as rehearsal may instead be attributed to a multi-modal refreshing mechanism (Cohen et al., 1997; M. R. Johnson et al., 2015). Indeed, there is evidence that suppressing an articulatory rehearsal process (e.g. via repetition of the word '*the*') will cause participants to recruit additional, domain-specific, maintenance resources (Chein & Fiez, 2010; Gruber & Cramon, 2003). Consequently, these control mechanisms may accommodate n-back performance for non-verbal stimuli such as fractals, abstract shapes (Nystrom et al., 2000; Ragland et al., 2002), faces (Dade et al., 2001), and olfactory stimuli (Dade et al., 2001; Jönsson et al., 2011).

Further distinction between n-back strategies considers the use of *control.* That is, whilst rehearsal strategies reflect *high control*, other strategies modelled for the task propose *low control* processes that allow effective performance without rehearsal of the memoranda (e.g. Juvina & Taatgen, 2007; B McElree, 2001). For example, a slow, reactive search process has been proposed to aid recovery of order information in circumstances where the n-back item is not maintained in focal attention (McElree, 2001). In this model, the most recent item is retrieved, which cues the next in the sequence, and so on until a match to the probe item is found. Alternatively, Juvina and Taatgen (2007) propose a low control strategy that compares a temporal-based estimation of a target item's encoding (a 'time-tag') to the approximate age of a target item (see also Nijboer, Borst, van Rijn, & Taatgen, 2016). To be clear, rather than actively maintaining the *n* items in memory, participants rely upon residual memory traces of the items to judge whether the current item matches the activation level expected of a target (i.e. a temporal familiarity judgement).

A variable task demand that may affect the employment of a particular n-back strategy is the prevalence of recent lures (Ralph, 2014). In his thesis, Ralph (2014) assessed whether the ratio of targets to lures influences the control strategies adopted in the n-back task, hypothesising that a lure-heavy sequence will increase the use of proactive control (i.e. engaging attention to the stored target item prior to presentation of the probe item). In direct contrast to the prediction, fewer target trials in a sequence resulted in a decrease in the use of a proactive control method. In explanation, he proposes the effort involved in keeping targets active is not worth the reward of getting relatively few trials correct when the number of targets is low. Consequently, his findings provide evidence that participants are able to extract information about the nature of the task, and to make adjustments to their control strategy accordingly (Ralph, 2014).

It is also possible that participants may make a strategic decision to make n-back judgments based only on familiarity (Juvina & Taatgen, 2007). Above-chance performance during the n-back task is possible if a strategy is adopted to accept probes when a familiarity-strength criterion is exceeded, and this method is likely the primary process used to reject non-recent lure items (Harbison et al., 2011). However, the inclusion of recent lures mean familiarity cannot be used to accurately identify a target item, and a cognitive control process must be adopted to discriminate these lures from targets to prevent inflated false alarm rates (Juvina & Taatgen, 2007; Szmalec et al., 2011). It appears that participants will typically opt for control strategies (i.e. a strategy that attempts to explicitly compare the probe item to information linked to the n-back serial position) to maximise accuracy (see McElree, 2001), though sequences with a low number of recent lure items can increase reliance on familiarity (Harbison et al., 2011). This strategy may also occur if there are failures in recollection that prevent a judgement of the probe's position (Juvina & Taatgen, 2007). Recollection-based decisions must be made to ensure a target or recent-lure decision is correct, and it is this process that is proposed to engage cognitive control processes (Juvina & Taatgen, 2007; Smith & Jonides, 1999).

How cognitive control strategies in the n-back task may be applied to an olfactory n-back task is unclear. Verbal recoding would allow the use of proactive or reactive high-control rehearsal strategy (Chatham et al., 2011; Juvina & Taatgen, 2007), and could explain a relationship between odour verbalisability and n-back performance (Jönsson et al., 2011). However, the dissociation of verbal and olfactory short-term memory indicates that a purely verbal rehearsal process may not be appropriate (Andrade & Donaldson, 2007). A high-control strategy may instead be applied to an olfactory representation through a refreshing process (e.g. M. R. Johnson et al., 2015). This refreshing process involves directing attention to one of several active representations (M. R. Johnson et al., 2015; Raye et al., 2007), and is proposed to drive non-verbal rehearsal (Baddeley, 2012). However, olfactory imagery has been proposed unique in its inability to give rise to a conscious olfactory representation (Stevenson & Attuquayefio, 2013; Zucco, 2003), meaning such rehearsal-based strategies may not be possible. If this is the case, the strategy might be mediated by whether an odour is identified or named. For hard-to-name odours, this may involve the adoption of a low-control strategy, which involves a reactive memory search, or a familiarity-based temporal estimation (Juvina & Taatgen, 2007; B McElree, 2001). Alternatively, failure to recollect these low verbalisability odours may result in the adoption of a familiarity that does not include any control process (i.e. acceptance of an item if a familiarity strength signal exceeds a fixed criterion, Juvina & Taatgen, 2007).

3.0.4 Olfactory learning and the role of familiarity

As noted above, the characteristics of the odours may be important in the type of strategy adopted in the olfactory n-back, and, consistent with past work (Jönsson et al., 2011), the present study categorises odours according to a measure of verbalisability scores. However, familiarity scores correlate strongly with verbalisation scores (r = .84, see Chapter 2). High odour familiarity is related to processes of perceptual learning, which can shape future perception of items (Goldstone, 1998). Consequently, it is important to consider the effects of both normative familiarity and experimental familiarity on working memory performance in the n-back task.

Perceptual learning is an integral part of the object recognition account of olfactory processing (Stevenson & Boakes, 2003; Wilson & Stevenson, 2006). That is, olfactory perception is proposed to involve matching an input pattern from receptor activation to previous encodings contained within an object store (Stevenson & Wilson, 2007). A novel or unfamiliar odour will activate many stored objects and result in a vague representation, whilst representations of familiar odours are more stable due to strong activation of only a few stored encodings (Stevenson & Mahmut, 2013). Consequently, two unfamiliar odours will be less discriminable from one another and judged to have more similarities than two familiar odours, due to both of them being redolent of many other odours (Mingo & Stevenson, 2007; Stevenson, 2012).

The importance of odour perceptual learning on perception has been shown in both experimental and naturalistic settings. In naturalistic studies, general olfactory perceptual experience can mediate discriminability of odours (Stevenson & Boakes, 2003). For example, wine experts and regular wine drinkers have both shown increased discriminability of odours compared to non-experts. This finding indicates an advantage from increased exposure can occur through perceptual learning, and that because experts and regular drinkers saw similar improvements the effect does not appear to be due to increased knowledge of the stimuli (Melcher & Schooler, 1996).

Experimental manipulation of familiarity through repeated exposures can also improve discriminability of odours (Li, Luxenberg, Parrish, & Gottfried, 2006; Rabin, 1988; Stevenson, 2001; Wilson & Stevenson, 2003b), and improve both short and long-term memory (Jehl et al., 1995, c.f. 1997; Nguyen et al., 2012; Valentin et al., 2011). Importantly, these improvements in recognition performance have occurred when odours were previously unfamiliar and the task did not include instructions to label the odours (Jehl et al., 1995). Though this supports effects of perceptually-based familiarisation, it should be noted that improvements in Jehl et al. (1995) were only observed on the proportion of false alarm rates. Furthermore, in a follow-up study of long and short term odour memory, Jehl, Royet, and Holley (1997) found no effect of familiarisation on recognition memory unless accompanied with learned verbal labels. Indeed, the experimental familiarisation condition even saw a disruptive effect on long-term recognition memory, which they attribute to confusion in the chronology of the pre-exposure and testing stages. More recently, Nguyen, Ober, and Shenaut (2012) showed an improvement in recognition performance (*d'*) when targets were experienced in multiple encoding trials. A trend was also observed for this improvement to be greater for less familiar odours, though the interaction was non-significant. They suggest multiple exposures can improve recognition performance by increasing the distinctiveness of items at encoding.

Though there is some support for a perceptual learning effect that can influence the representation of odours in memory, these effects may be confounded or complemented by verbal memory (Stevenson, 2001; Stevenson & Boakes, 2003). Working memory capacity is typically defined in terms of the number of *chunks* that can be maintained (Cowan, 2001), and that these chunks are organised by learned information in long-term memory (e.g. Hulme, Maughan, & Brown, 1991; Thorn, Gathercole, & Frankish, 2002). Consequently, increased knowledge of the processed stimuli can improve working memory capacity by facilitating the formation of these chunks (Jackson & Raymond, 2008). Indeed, novel visual stimuli with no verbal or semantic associations may result in a capacity of only one item (H. Olsson & Poom, 2005). Furthermore, whilst Melcher and Schooler (1996) suggest an expert discriminability advantage that is due to exposure alone, comparisons of olfactory short-term memory in wine-tasters to trained panellists and non-experts has shown an expert advantage that is better attributed to an ability to verbalise the odorants (Valentin et al., 2011). These semantic memory effects are proposed to occur as a consequence of effortful association of labels to the odorants (Stevenson, 2001), and odour labelling is likely to occur in an explicit working memory task such as the n-back (Jönsson et al., 2011). Consequently, verbal learning may be responsible for any discriminability and working memory advantages for familiar odours, or may improve odour discrimination over and above the effects of mere exposure alone (Rabin, 1988; Stevenson & Boakes, 2003).

3.1 Experiment 1: Olfactory n-back partial replication

3.1.1 Introduction

Experiment 1 is a partial replication of Jönsson et al. (2011). The purpose of this replication is threefold. First, above chance olfactory n-back performance has only been shown in two previous studies (Dade et al., 2001; Jönsson et al., 2011), so the experiment seeks to replicate this finding with a different set of odours. Second, this experiment can validate recent normative data from Chapter 2, by demonstrating the same facilitative effect for highly verbalisable odours as that reported by Jönsson et al. (2011). Indeed, the findings in Chapter 2 indicate that verbalisability of odours is an unsuitable dimension with which to control odours, due to high levels of individual differences. However, not only has a similar measure been used successfully in other research (Jönsson et al., 2011), but the verbalisability scores in Chapter 2 correlated strongly with other dimensions deemed suitable for use in olfactory memory experiments. Although this is not tested, one might speculate that verbalisability for low and high extremes of the verbalisability dimension are less susceptible to individual differences, and consequently the use of these odours enable an effective manipulation of odour verbalisability in the present tasks. Third, an additional testing sequence is introduced to investigate discriminability changes due to perceptual and verbal learning throughout the task (Jönsson et al., 2011, employed a single testing sequence). That is, does performance improve for low verbalisability odours as a result of repeated exposure to those odours?

The present experiments use an index of recognition ability derived from signal detection theory, and additionally reports independent analyses of hits and false alarms. This is necessary to provide insight into specific changes in the ability to reject lures and accept targets, in addition to shifts in response bias. For example, the findings in Lyman and McDaniel (1986) have been criticised due to the facilitation from labelling on recognition performance occurring through a reduction in false alarms rather than an increase in target recognition. However, signal detection is preferred over hit rates because nameability of an odour has been proposed to affect response strategy, with an item that is not identified judged as 'new' more frequently (R. A. Frank et al., 2011). This may be because the information about an odour that is not named is very limited (Jönsson & Olsson, 2003), with participants therefore reluctant to respond *old* when they cannot report any information about the odorant.

Three hypotheses are presented based upon previous evidence of a working memory advantage for verbalisable odours and considering the effect multiple exposures may have on n-back strategy. It is predicted that (1) odour working memory will be above chance for low verbalisability odours but (2) performance for high verbalisability odours will be significantly better (Jönsson et al., 2011). Across testing sequences, it is predicted that (3) greater improvement for low verbalisability odours from repeated presentations, through a process of perceptual learning and verbal learning (Nguyen et al., 2012; Stevenson, 2001; Stevenson & Mahmut, 2013a).

3.1.2 Method

3.1.2.1 Participants

Twenty participants (12 males and 8 females, mean age = 20.0, *SD* = 2.7) participated in exchange for course credit. Participants who self-reported olfactory impairments (e.g. symptoms of cold) and smoking (Katotomichelakis et al., 2007) were excluded, as were participants aged over 40 years (Doty et al., 1984). Ethical approval was obtained via the Bournemouth University Ethics Committee.

3.1.2.2 Materials

The odours were as described for Chapter 2.

Twelve odours were randomly selected from the twenty highest and lowest verbalisability scores to form the low and high verbalisability odour sets used in the n-back task (see Appendix B). The verbalisability judgment in Chapter 2 followed closely that of Jönsson et al. (2011) such that stimuli were scored from 0-3 according to the quality of the verbal labels provided, with a lower score indicating vague or absent verbalisability and a higher score reflecting use of a specific noun. Verbalisability for the two odour sets differed significantly, $t(12) = 26.38$, $p < .0005$, $d = 15.23$, $BF_{10} >$ 1,000 ($M_{\text{high}} = 2.66$, $SD_{\text{high}} = 0.11$; $M_{\text{low}} = 1.12$, $SD_{\text{low}} = 0.09$). An additional two odours were selected from the high and low verbalisability odorant samples. These were chosen to act as non-analysed buffer items (i.e. these are used at the start of the task and are not included in the analysis).

As discussed in Chapter 2, a normative verbalisability score may not be the most suitable dimension on which to base odour selection, due to high variability across participants. However, familiarity scores deemed more suitable for odour selection covaried with verbalisable ratings, such that the odour sets were also significantly different across familiarity scores, $t(12) = 21.62$, $p < .001$, $d = 11.57$, $BF_{10} > 100$.

Eight line drawings, printed on individual A5 sheets of paper, were taken from Snodgrass and Vanderwart (1980) and used as a 2-back practice task at the start of the experiment.

3.1.2.3 Design

A continuous yes/no recognition task was employed on two testing sequences of 52 odour trials, where each trial necessitated a judgment as to whether the present odour was the same or different to the odour presented two items previously (i.e. the 2-back task). The experiment employed a within-participants multifactorial (2x2) design. The first within-participants factor concerned whether the block of odour trials contained odours categorised as high or low on verbalisability. This was operationalised as a block of 26-trials employing high verbalisable odours and a block of 26-trials employing low verbalisable odours. There was no interval between blocks (i.e. it was presented as a continuous 52-trial sequence). The second within-participants factor concerned testing sequence. Participants undertook two 52-trial testing sequences, with each testing sequence containing a block of high and low verbalisable odour trials. These odours were the same items used in the first sequence. The presentation order of trials was predetermined before testing, and the order of blocks was counterbalanced via a Latin square design.

Within each (high or low verbalisability) block, the six (high or low verbalisable) odorants appeared as a ‘target’ once (25% of trials), and three times as a ‘lure’ (75% of trials). Targets were odorants that had been presented two trials previously, and thus required a ‘yes’ response. Lures were odorants not matching the odour presented two trials previously, and therefore required a 'no' response. Thus, a block comprised 24 critical trials (and 2 buffer trials) with each odorant presented four times. The first two trials in a sequence would always be lures, so preceding each block were two additional buffer trials. For the high verbalisability block, 'Pear' was presented for these two trials, and for the low verbalisable block 'Nag Champa' was presented. These buffer odours were not repeated elsewhere in the sequence, and responses for the buffer trials were not entered into the analysis. Recent lures at positions $n+1$ and $n-1$ were allowed to occur, and randomly appeared in sequences. Differences in the target to recent lure ratio across participant trials were equated across verbalisability conditions using the counterbalancing methods described below.

When determining the order of trials within blocks, the nature of the 2-back task required that six lures were tethered two positions before the six matching targets. To be clear, for that odour to be a target, it must first be employed as a lure two trials previous. The remaining 12 lures in each block were placed pseudo-randomly, with the caveat that their position did not result in itself or a previously positioned lure becoming an unintended target, nor result in a target becoming a lure. The predetermined trial orders were counterbalanced, such that a sequence of lures and targets was re-used for another participant with the alternative set of 6 odours.

The number of correct target identifications (Hits), and incorrect identifications of a lure as a target (False Alarms, FA) were recorded and used to compute the proportion of Hits to FA via *A'*. The mismatched number of recent lures across participants made analysis of only these lure types for incidences of false alarms unsuitable. Instead, false alarms were calculated from all lure probes, at the cost of having slightly inflated correct rejection proportions and *A'* scores. This measure of signal detection theory was selected due to the unequal trial numbers for lure and targets, and because it allows FA rates to exceed Hits. *A'* was calculated as 0.5 + ((Hits – FA) x (1 + Hits – FA)) / ((4 x Hits) x (1 – FA)) when Hits exceeded FA, and as 0.5 – ((FA – Hits) x (1 + FA – Hits)) / ((4 x FA) x (1 – Hits)) when FA exceeded Hits (Stanislaw & Todorov, 1999). Unlike *d'* where Hit rates of one or FA rates of zero result in an indefinite value, use of *A'* allows these results to remain unadjusted.

3.1.2.4 Procedure

The experiment was conducted in a quiet, well-ventilated room with a fan to circulate fresh air. Participants sat opposite the experimenter, separated by a wooden screen with a central fixation cross to prevent visual inspection of the odorants. Prior to the olfactory task, participants performed an 8-item visual version of the 2-back task in order to familiarise themselves with the procedure.

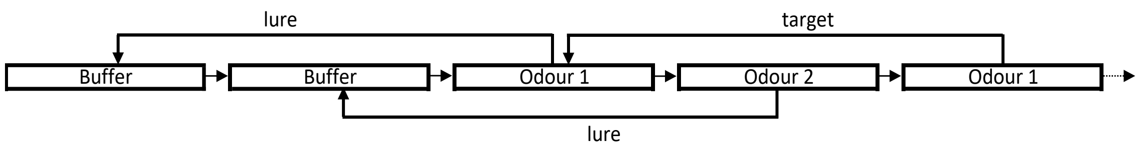

Figure 2. Schematic diagram of the 2-back task. Two buffer items precede the 24 test trials.

The 2-back task (Figure 2) presents participants with a sequence of stimuli, where each item must be compared with the stimulus presented 2 trials previously, whilst simultaneously remembered for comparison in future trials. Participants completed two sequences of 52 trials, where each trial consisted of a single odour presented under the nose of the participant for 2 seconds. Participants were required to make a verbal 'yes' response if the currently presented item matched the odour presented two trials previously and a 'no' response if it did not. An 8-second inter-stimulus interval (ISI) separated odour presentations. In the interval between the two 52-odour sequences, participants were encouraged to drink water and given a 5-minute break. Total testing time (including breaks) took approximately half an hour.

3.1.3 Results

Data in this and subsequent experiments were analysed using traditional analysis of variance (ANOVA) and planned Bonferroni-corrected comparisons. However, in Chapter 3 and Chapter 4, Bayesian ANOVA with default priors were also performed using JASP (Love et al., 2015; R. D. Morey & Rouder, 2015; Rouder, Morey, Speckman, & Province, 2012). This is a model-based approach, where models containing main effects and interactions can be compared. To be clear, this process produces a *Bayes Factor* value that indicates the ratio of support for one model over another. Typically, the comparison will be between a particular model (for example, a model with both main effects) and the null model, producing a Bayes Factor indicating the level of support for this model. However, Bayes Factors are transitive, and the assessment of models additive, so a model with an interaction term added can also be compared to this main-effects model. This produces a Bayes Factor indicating the strength of evidence for an improvement to the model when the interaction term is included. A typical cut-off as providing substantial support for a model is for a likelihood given the data three times greater than the likelihood for the null (Jeffreys, 1998). This equates to a Bayes Factor greater than 3 as substantial support for the alternative hypothesis, and below 1/3 as substantial support for the null. A score between 1/3 and 3 indicates insensitivity to either hypothesis.

Bayes Factors are also calculated for paired comparisons, which using the same cut-offs above outlines the strength of evidence for or against an alternative hypothesis. For Chapter 3 and 4, these use a default Cauchi prior distribution (Rouder et al., 2012). These are presented with *p* values and t-test results where appropriate, in the format

BF_{10} when testing the alternative hypothesis against the null, and BF_{01} when testing the null hypothesis against the alternative hypothesis. That is, $BF_{10} < 0.33$ and $BF_{01} > 3$ would indicate identical support for the null hypothesis.

3.1.3.1 A' sensitivity

Figure 3(A) reports *A'* for the high and low verbalisability groups, across testing sequences. An ANOVA was performed where the first factor was testing sequence (first and second) and the second factor was odorant verbalisability (low and high). The main effect of testing sequence was non-significant, $F(1, 19) = 0.35$, $p = .560$, $\eta_p^2 = .02$, indicating no overall change in recognition sensitivity over sequences. A significant main effect of verbalisability was found, $F(1, 19) = 5.95$, $p = .025$, $\eta_p^2 = .24$, with greater sensitivity for the high verbalisability odours ($M = .84$, $SEM = .01$) compared to low verbalisability odours ($M = .79$, $SEM = .02$). Importantly, the interaction between sequence number and verbalisability was significant, $F(1, 19) = 8.32$, $p = .010$, $\eta_p^2 = .31$. Bayesian ANOVA indicated strongest support for a model with main effects and an interaction between verbalisability and sequence ($BF = 4.24$ vs the null model), preferring this model over a main effects model by a factor of 8.08.

Follow-on Bonferroni-corrected paired comparisons ($\alpha = .025$) and Bayes Factor analysis revealed evidence against lower sensitivity for the low verbalisability odours ($M = .81$ $SD = .10$) compared to the high verbalisability odours ($M = .80$, $SD = .09$) in the first testing sequence, $t(19) = .41$, $p = .690$, $d = .12$, $BF_{10} = 0.25$. However, the second sequence saw strong evidence for lower hits for low verbalisability odours ($M = .77$, $SD = .16$) compared to high verbalisability odours ($M = .88$, $SD = .07$), $t(19) = -3.58$, $p = .002$, $d = -.92$, $BF_{10} = 40.40$. That is, an effect of greater sensitivity for highly verbalisable odours was present only in the second testing sequence. There was evidence for a difference across testing sequences for high verbalisability odours that is suggested to be driving this interaction, where performance was better in the second testing phase compared to the first, $t(19) = -2.92$, $p = .009$, $d = -.91$, $BF_{10} = 5.69$. In comparison, there was anecdotal evidence against a difference between performance for the low verbalisable odours in both testing sequence, $t(19) = 1.27$, $p = .220$, $d = .36$, $BF_{10} = 0.47$.

Using a single sample t-test, *A'* sensitivity scores were also analysed against a chance score of 0.5. There was strong evidence for above chance performance for the low verbalisability odours in both the first, $t(19) = 13.60$, $p < .001$, $d = 6.24$, $BF_{10} > 1{,}000$, and second sequences, $t(19) = 7.63$, $p < .001$, $d = 3.50$, $BF_{10} > 1{,}000$. To be clear, this demonstrates that in both testing sequences, sensitivity was above chance for the low verbalisable odours.

3.1.3.2 Hit rate analysis

A 2 (sequence: first, second) x 2 (verbalisability: low, high) ANOVA was performed for hit rates, calculated from correct target recognition. Figure 3(B) shows these hit rates across odorant verbalisability and sequence number. The analysis revealed a non-significant main effect of testing sequence, $F(1, 19) = .49$, $p = .494$, $\eta_p^2 = .03$. However, there was a significant main effect of odorant verbalisability, $F(1, 19) = 5.94$, $p = .025$, $\eta_p^2 = .24$, with greater hits for odours high on verbalisability ($M = .67$, $SEM = .03$) than low on verbalisability ($M = .58$, $SEM = .04$). This effect of verbalisability interacted across testing sequence, $F(1, 19) = 5.15$, $p = .035$, $\eta_p^2 = .21$. However, Bayes Factors were insensitive to any preference for a main effects model or interaction model over the null.

Analysis of the interaction using Bonferroni-corrected paired comparisons ($\alpha = .025$) and Bayes Factors revealed evidence against greater hits for verbalisable odours compared to low verbalisability odours in the first testing sequence, $t(19) = -.38$, $p = .705$, $d = .10$, $BF_{10} = 0.32$, but evidence for greater hits for verbalisable odours ($M = .72$, $SD = .18$) compared low verbalisability odours ($M = .57$, $SD = .26$) in the second testing phase, $t(19) = -3.21$, $p = .005$, $d = .67$, $BF_{10} = 19.73$. In summary, the findings replicate those shown for *A'* sensitivity, where a target recognition advantage was present for high verbalisability odours in the second sequence.

3.1.3.3 False alarm rate analysis

Figure 3(C) shows the false alarm rate for low and high verbalisability odours across testing sequences. The analysis revealed a non-significant main effect of sequence, $F(1, 19) = 1.22$, $p = .283$, $\eta_p^2 = .06$, and differ from the results observed for hits and *A'* with also a non-significant main effect of verbalisability, $F(1, 19) = .50$, $p = .487$, $\eta_p^2 = .03$. There was however, an interaction between testing sequence and verbalisability, $F(1, 19) = 5.98$, $p = .024$, $\eta_p^2 = .24$. Indeed, an interaction-only model was preferred to the null by a ($BF = 7.11$ vs. a null model).

Bonferroni-corrected paired comparisons ($\alpha = .025$) and Bayes Factors revealed evidence against greater false alarms for low verbalisability odours in sequence 1, $t(19) = -1.99$, $p = .061$, $d = -.46$, $BF_{10} = 0.09$, but evidence for greater false alarms in the second sequence for low verbalisability odours ($M = .17$, $SD = .10$) compared to verbalisable odours ($M = .11$, $SD = .09$), $t(19) = 2.14$, $p = .046$, $d = .58$, $BF_{10} = 2.90$. In summary, the findings showed a significant interaction similar to those observed for hits and *A'*, but without a main effect of verbalisability. In summary, these data provide some evidence for lower false alarms for verbalisable odours in the second testing sequence.

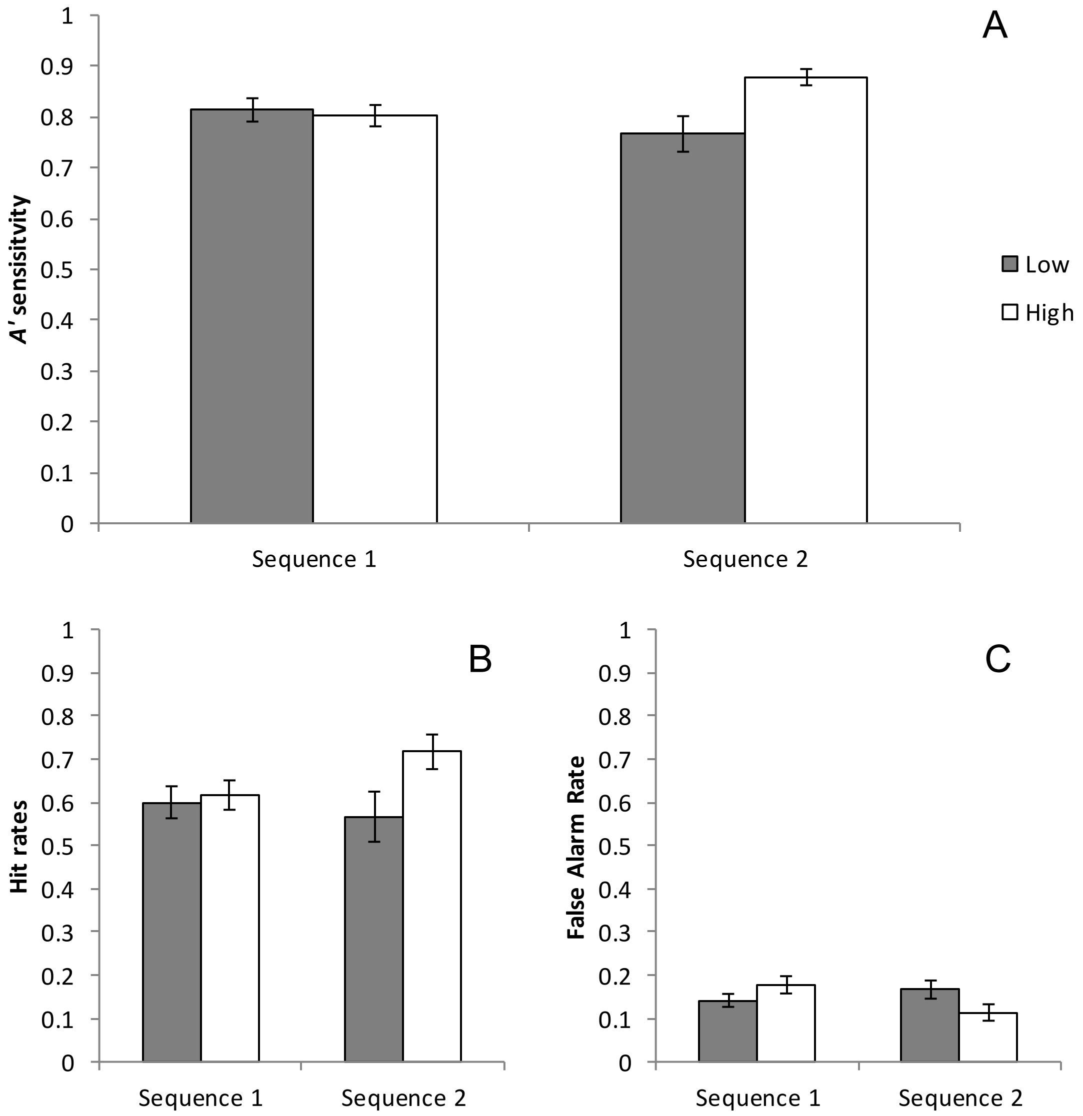

Figure 3. (A) *A'* sensitivity, (B) false alarm rates, and (C) hit rates, for low and high verbalisability odours, across testing sequences. Error bars denote 1 standard error of the mean.

3.1.4 Discussion

Experiment 1 examined the extent to which olfactory working memory performance is affected by the verbalisability of the odours. Similar to the findings of Jönsson et al. (2011), superior n-back sensitivity is reported for odours classified as exhibiting high verbalisability, whilst still showing above chance performance for odours classified as exhibiting low verbalisability ($p < .005$, $BF_{10} > 1{,}000$, for low verbalisability *A'* in a one-sample t-test against the chance score of 0.5). This replication of Jönsson et al. (2011) provides further validation for the normative ratings reported in Chapter 2 in respect to ratings of verbalisability (see also Chapter 5 for differential effects of verbalisability on proactive interference).

The present findings differ to Jönsson et al. (2011), however, in reporting differences between high and low verbalisable odours in the second sequence only. The working hypothesis for this experiment was that differences between high and low verbalisable odours may attenuate in the second sequence due to a refinement/development of functional labels for the low verbalisable odours following repeated exposure. In contrast, the benefit for the high verbalisable odours only emerged in the second sequence. This trend was mostly consistent across *A'*, hit rates, and false alarms.

The gradual development of verbalisability effects is perplexing; whilst, differentiation between odorants is shown to improve through repeated or prolonged exposure (e.g. Li et al., 2006; Rabin, 1988), these effects should not be confined to a specific increase in discriminability for high verbalisability odours only. It is possible an asymmetry from the effect of multiple exposures occurred due to an inability to effectively perform the task with low verbalisability odours. However, performance for these odorants was above chance, and no participants reported an inability to smell any odorant. This interaction across testing sequences therefore raises some questions over the use of verbal codes, mere exposure effects, and the development of strategies throughout the n-back task. It is possible, for example, that verbalisable odours have these labels refined throughout the task, resulting in improved n-back performance.

An odour representation in memory has been proposed to include both perceptual and verbal information, though the relative weight of these information types may vary as a function of being identified, or through the quality of available verbal information (e.g. Yeshurun et al., 2008; Zelano et al., 2009). A possible explanation is that initial poor quality verbal labels applied to the verbalisable odours were refined following multiple exposures (see Stevenson, 2001), or that some strategy shift occurred across testing sequence that favoured the high verbalisability odours. Indeed, the importance of label consistency has been stressed for the facilitative effect of verbal labelling in olfactory memory (R. A. Frank et al., 2011; Kärnekull et al., 2015), and labelling consistency is likely to increase with familiarity (R. A. Frank et al., 2011). Furthermore, there was a trend towards poorer performance for low verbalisability odours in the second sequence, which would be predicted if a verbal strategy was across the trials and applied to odours with a poor-quality label. Consequently, verbal learning can potentially explain the improvement for these odours as a result of an improved ability to categorise and label the stimuli.

3.1.4.1 Discriminability improvements

This section considers whether the advantage for high verbalisability odours in the second sequence was due to perceptual or verbal learning. Although the manipulation of odours was across a verbalisability dimension, the relationship between odour verbalisability scores and familiarity is high, such that verbalisable odours had high normative familiarity. However, high normative familiarity would predict a discriminability advantage for odours in the first testing sequence (Wilson & Stevenson, 2006), and this was not observed. Furthermore, olfactory perceptual learning would not predict a specific increase in discriminability for only high verbalisability odours. Indeed, the object recognition model of olfaction describes rapid perceptual learning that would predict either generalised improvements for both odour sets (Li et al., 2006; Rabin, 1988; Stevenson, 2001), or lower perceptual learning in the high verbalisability odour set because of these items' already-high normative familiarity (Goldstone, 1998; Nguyen et al., 2012; Stevenson, 2001; Stevenson & Mahmut, 2013a). Consequently, the findings indicate that the working memory advantage for high verbalisability odours was not due to perceptual learning of odours, either through previous pre-experimental normative familiarity or from experimental exposure-based familiarity.

An alternative consideration is the effect on odour discriminability from verbal learning (Stevenson, 2001; Stevenson & Boakes, 2003). That is, effortful association of verbal labels to the olfactory stimulus can improve item discriminability (Rabin, 1988; Stevenson, 2001), where retrieval of a label facilitates the matching process between perceptual input and stored representation (Wilson & Stevenson, 2006). Furthermore, label consistency has been stressed for the facilitative effect of labelling in olfactory memory (R. A. Frank et al., 2011; Kärnekull et al., 2015), and labelling consistency is likely to increase with familiarity (R. A. Frank et al., 2011). However, odour verbalisability, like familiarity, is a normative property of the stimuli that would predict better n-back performance due to a discriminability advantage in the first testing sequence (Jönsson et al., 2011; Lyman & McDaniel, 1986). The absence of such an advantage in the first sequence may be explained by poor quality verbal labels, which were then refined over multiple exposures to the odour (Stevenson, 2001). Consequently, a consistently applied, high quality label may enable a strong, stable representation in memory by nature of a strongly activated odour object.

Some aspects of the methodology further support a role of verbal learning. Although participants were aware of the number of trials they would perform, they were given no indication of the number of odorants used, nor had they smelled any of the test odorants before the task had begun. The first sequence is therefore characterised by the participant's lack of knowledge about the number of odorants and of the similarities between them. Consequently, the specificity of verbal label required to make an accurate distinction between each stimulus was unknown in the initial testing sequence. These findings are not consistent with Experiment 1 of Jönsson et al. (2011, p. 1026), where participants performed a mixed-blocks version of the n-back task with no prior exposure to odours other than a ‘few’ practice trials. In that experiment, Jönsson et al. showed a verbalisability advantage in their first and only sequence, and it seems unlikely that refinement of verbal labels would have occurred. However, in Experiment 2 of their study (which most closely matches the present procedure and adopts a blocked-verbalisability design), participants performed a discriminability task before completion of the n-back procedure. This amounted to 6 presentations of each odour before the n-back task was completed, and thus provided the verbal learning opportunities required to match the finding observed in this experiment.

3.1.4.2 N-back strategy shifts

A possible explanation for these data is therefore that high verbalisability odour n-back performance improved as participants refined and made more consistent the initial label attributed to each odour. It is important to note, however, that in Jönsson et al.’s (2011) assessment of discriminability for high and low verbalisable odours they concluded that discriminability advantages alone cannot explain the advantage observed. That is, participants first performed a discriminability test where each task odorant was paired in comparison trials with it and all other odours within the set. An *A’* score was calculated from the discriminability test and was directly compared to *A’* performance in the n-back procedure. The working memory advantage for high verbalisability odours was greater in magnitude than the differences in item discriminability, indicating an advantage for verbalisable odours in working memory that goes beyond discriminability (though Jönsson et al. accept that there may be problems with equating the differences across the two measures).

Consequently, although odour discriminability is a limiting factor in olfactory n-back performance, the use of verbal labels within an olfactory representation may also reflect performance differences based on underlying mnemonic strategies. First, however, whether the present findings might be accommodated within a high-control verbal rehearsal strategy is discussed (e.g. Chatham et al., 2011; Juvina & Taatgen, 2007). That is, the proposed role of verbal codes in odour working memory may support the use of a rehearsal strategy to update item information during the n-back task (e.g. based on the similar processes proposed for remembering named odours and verbal information, Olsson, Lundgren, Soares, & Johansson, 2009). In Jönsson et al. (2011), they suggested that low verbalisability odorants elicit at least some spontaneous verbal association. The present study's initial assessment of verbalisability, where a score of zero included hedonic responses (e.g. 'disgusting'), would presumably also allow a weak level of verbal coding. This weak label may have provided the appropriate code for a rehearsal strategy, and be sufficient for above chance recognition performance. Indeed, the quality of verbal label for an odour may be directly related to the ability to effectively rehearse that odour, leading to the advantage for the high verbalisability odours.

However, whilst these findings indicate that verbal labels were refined over time, this only occurred for verbalisable odours. This indicates some difference in the way low verbalisability odours are represented. Indeed, there is evidence for some shift in the representation of olfactory information based on whether odours are identified (e.g. Zelano et al., 2009), or when odours are not identified that a perceptually-based recognition may occur, known as *recognition without identification* (Cleary et al., 2010).

If these low verbalisable odours are represented by a perceptual code, attentional refreshing (M. R. Johnson et al., 2015; Raye et al., 2007) may be used to allow a high-control rehearsal strategy without the need for verbal recoding (Juvina & Taatgen, 2007). However, it is also possible that the nature of these perceptually represented odours make such strategies unsuitable (Stevenson, 2009), and these odours might instead be dealt with using an alternative n-back strategy. That is, the availability of verbal information, and resultant changes in the way odours are represented in memory, may affect the way in which the rehearsal window is maintained and the method for making a comparison between n^{th} and trial item (Juvina & Taatgen, 2007).

A plausible strategy for low verbalisability odorants is as a *low-control* 'time tag' strategy (Juvina & Taatgen, 2007). To be clear, this strategy compares the activation strength (familiarity) of a probe item to a stored estimate of activity for a target item. Adoption of this strategy for odours can explain the poorer working memory performance for these low verbalisability odours, as the strategy is proposed to be a noisy method of estimating target appearances (Juvina & Taatgen, 2007). A familiarity signal is a relatively automatic source of information for recognition (Loaiza, Rhodes, Camos, & McCabe, 2015), though some controlled processing is required for the comparison of this signal to the estimate of target signal strength. However, a familiarity-based strategy may alternatively accept an item if the strength signal falls above a certain criterion. This strategy therefore differs to the time-tag strategy in the amount of control applied to interpret this familiarity signal. This strategy can be sufficient for demonstrating above-chance performance for the low verbalisability odours, as a decision to reject most lures, and to accept targets, can be made based on a familiarity signal alone. However, the inclusion of close lures, where its previous presentation falls close to the n-back position, means participants are more likely to make false alarms if relying on a familiarity-based strategy (Kane et al., 2007). Whilst the present findings did show poorer *A'* sensitivity scores for low verbalisability odours that were due in part to greater false alarms, a reliance on a familiarity criterion alone would predict a corresponding increase in hits due to acceptance of any recently presented item, and this was not observed.

In summary, Experiment 1 has shown that working memory ability for odours is improved when odours are classified as highly verbalisable, compared to those odours classified as hard-to-verbalise. This working memory advantage for highly verbalisable odours occurs only after experiencing the odours in an initial n-back task, indicating learning related to the odour representations following exposure, perhaps through refinement of a verbal code. Importantly, there was little improvement across sequences for low verbalisability odorants. This may reflect a differing reliance on olfactory and verbal codes between the two odour sets, rather than a linear improvement due to the quality of verbal code used for each odour.

3.2 Experiment 2: Assessment of maintenance strategies with dual tasking

3.2.1 Introduction

Experiment 1 replicated the superior n-back performance for nameable odours reported by Jönsson et al. (2011). One might prosaically interpret this finding as an employment of verbal labelling for the odours thereby enabling the employment of a high control n-back strategy incorporating rehearsal (e.g. Harbison et al., 2011; Juvina & Taatgen, 2007; Szmalec et al., 2011). Moreover, given that the difference between nameable and non-nameable odours was only found in the second sequence of odours, it suggests that the labels assigned to the verbalisable odours and/or verbal rehearsal strategy may become more refined over repeated exposures.

Experiment 2 tests the assumption that the n-back benefit for verbalisable odours derives from verbal labelling/rehearsal through the employment of a dual-tasking procedure, specifically a secondary verbal task. This experiment applies a dual-task paradigm to examine the strategies used in an olfactory n-back procedure. In a dual-task procedure, tasks that are thought to occupy the same or different processes are performed concurrently. A multicomponent working memory framework predicts interference from a secondary task if both tasks occupy the same sub-component, and little or no interference if the tasks occupy different systems (Baddeley, 1986; Logie, 2011). Concurrent articulation (CA) is a commonly used secondary task for suppressing the articulatory rehearsal process (e.g. Cocchini et al., 2002). The method has been shown to remove the word length effect (Baddeley et al., 1975) and phonological similarity effect (Saito et al., 2008) for visually presented words, suggesting that the conversion of stimuli into phonological representations is disrupted.

Concurrent secondary tasks have been used previously in olfactory research to examine the processes used to maintain olfactory information in memory. In a study examining the claim that olfactory recall and recognition is supported by a dual code, Annett and Leslie (1996) used secondary tasks designed to suppress verbal, visual, or both verbal and visual encodings. Participants were presented with 15 odours in an acquisition task, followed by an immediate recall or recognition task. During stimulus encoding, participants were required to verbally repeat digits as they were heard through headphones, track a character on-screen through complex mazes, or perform both tasks simultaneously. They found the combined task impaired performance more than both single tasks and the two single tasks impaired performance equally. They suggested a modification to the dual-coding account where multiple, independent, non-verbal systems exist to support a memory trace. However, these interference effects from multiple modalities may instead be related to task difficulty, and the taxing of executive resources. Indeed, it should be noted that a general executive resource/difficulty explanation means that one cannot falsify an amodal explanation using dual-tasking. In addition, it is also possible that effects of visuospatial interference occur due to verbal recoding of the visual task, which in turn may suppress a verbal memory trace (Annett & Leslie, 1996).

More recently, dual-tasking procedures have shown that a secondary verbal task may exhibit a limited effect on odour memory. Andrade and Donaldson (2007) investigated the effect of concurrent verbal, visual, and olfactory short term memory tasks on a primary verbal short term memory task (Experiment 1), and the effects of concurrent verbal, visual, and olfactory short term memory tasks on a primary olfactory memory task (Experiment 2). In Experiment 1, a verbal memory task was interfered by a secondary verbal task, whilst smaller interference effects attributed to a general resource load were observed for secondary olfactory and visual tasks. In Experiment 2, they demonstrated interference from a concurrent olfactory memory task on a primary olfactory task, and no effects from either verbal or visual secondary tasks. These findings are consistent with an independent memory system for olfactory stimuli that is not dependent on verbal coding. However, Andrade and Donaldson (2007) did not preclude a facilitative contribution of verbal labelling in certain olfactory tasks, stating that such labels can improve memory by providing an additional memory trace (Paivio, 1990). Furthermore, the extent to which this finding can be extrapolated to the n-back is questionable. First, the task demands of the n-back procedure may differ to that of a standard recognition task, and this may affect the requirement for verbal recoding. Second, it is not known to what extent the odours employed in Andrade and Donaldson (2007) were nameable. Specifically, whilst the corpus of 12 commercially available aromatherapy odours used by Andrade and Donaldson (2007) may have been familiar, it is unknown to what extent they could be verbalised (though one might expect verbalisability to be generally high). If these odours were hard-to-name at the outset, it would be unsurprising that recognition was unaffected by a manipulation that disrupts naming/verbal rehearsal (although see Miles & Hodder, 2005, for a main effect of CA on odour recognition).

Though the present experiment is the first to apply CA to an olfactory n-back procedure, distractor stimuli have been applied successfully in n-back inter-trial intervals with non-olfactory stimuli (Vuontela, Rämä, Raninen, Aronen, & Carlson, 1999). Participants performed 1- and 2-back tasks with location or colour memoranda, and were shown task irrelevant distractors consisting of different locations or colour stimuli in the trial intervals. Whilst only spatial distractors impaired spatial n-back performance, colour distractors only impaired visual n-back when verbal recoding was blocked by concurrent articulation. That is, concurrent articulation was effective in removing the verbal rehearsal of the n-back colours, meaning interference from the same visual modality was possible. This demonstrates that verbal strategies can be affected by CA during the n-back inter-stimulus-interval, thereby validating the manipulation in Experiment 2.

Also considered in Experiment 2 is that most n-back strategies involve several executive functions, including updating of items for maintenance, binding of items to serial position, and resolution of proactive interference (e.g. Chatham et al., 2011). Whilst the active maintenance of stimuli within modality-specific slave systems (Baddeley, 1986; Baddeley & Hitch, 1974) can be tested using dual-task procedures (e.g. Cocchini et al., 2002; Duff & Logie, 2001), the role of executive resources can also be examined by using tasks that are thought to occupy processes not engaged in the primary task's slave system. This was the approach in Simmons (2000), using more complex concurrent tasks than those used in Vuontela et al. (1999). Working memory tasks performed in the interval between 1 and 3-back trials included rhyme judgement, mental rotation, and random key presses. These tasks were proposed to load upon the phonological loop, visuo-spatial sketchpad, and central executive, respectively. Consequently, Simmons' hypothesis was that only verbal and key press tasks would impair verbal 3-back performance. However, they instead found little impairment from random key presses, whilst both rhyme judgement and mental rotation tasks had a disruptive effect. That is, the concurrent mental rotation task, thought to be associated with visuo-spatial memory, impaired verbal n-back performance. These findings may be explained by the increased use of executive resources during the mental rotation task (see Logie & Salway, 1990), though the pattern of interference may also fit a unitary resource model and explained by increasing task difficulty (Simmons, 2000). The present study therefore includes a concurrent rotation task to assess the effect of load on general resources in the olfactory n-back task, as a comparison to CA and control conditions.

Experiment 2 directly tests the extent to which the working memory benefit for high verbalisability odours is due to verbal labelling and/or rehearsal of those odours using CA, and the engagement of executive resources using concurrent mental rotation. The olfactory n-back methodology of Experiment 1 is replicated and includes concurrent secondary tasks in the inter-trial interval. The representation of unfamiliar, hard-to verbalise odours is proposed to rely on the olfactory perceptual code in working memory (e.g. Zelano et al., 2009), whereas a verbalisable odorant may make use of the additional label through some form of a dual representation (Paivio, 1990; Stevenson & Wilson, 2007; Yeshurun et al., 2008). By including CA during the inter-trial interval (a manipulation shown to disrupt verbal rehearsal in a visual n-back task, Vuontela et al., 1999), it attempts to disrupt the rehearsal of those labels and therefore should impair performance to a greater extent for the high verbalisable odours compared to the low. An additional aim in Experiment 2 is to replicate the unexpected finding from Experiment 1 that the advantage for high verbalisable odours is found only in the second sequence. If this effect is replicated, CA is predicted to be disruptive to the recognition of high verbalisability odours in the second sequence only. Finally, a secondary mental rotation task is predicted to employ executive resources (Simmons, 2000), and, as a consequence, it is predicted to exhibit a detrimental effect across the task and not interact with the verbalisability of the odours.

3.2.2 Method

3.2.2.1 Participants

Seventy-two participants (mean age = 19.82, SD = 2.93, 61 females, 11 males) participated in exchange for course credit. The same exclusion criteria as described for Experiment 1 were applied. None had participated in Experiment 1.

3.2.2.2 Materials

The olfactory stimuli were as described for Experiment 1.

The mental rotation task consisted of 104 unique images, showing horizontally presented pairs of three-dimensional objects in the style of Shepard and Metzler (1971) and obtained from Ganis and Kievit (2015). The left item formed the baseline object with which the right item, the target, was to be compared. Twenty-six unique baseline objects were used, and varied between 8 and 11 blocks long. The 52 congruent object conditions presented the target item identical to the baseline, rotated clockwise on a vertical axis 100, and rotated 150 degrees. The incongruent conditions used the same rotations of an object similar to the baseline object, but with a key difference such as a single arm direction pointing in the opposite direction (see Figure 4). Images were 800x427px and displayed in the centre of a 22-inch 60Hz monitor using stimulus presentation software OpenSesame (Mathôt, Schreij, & Theeuwes, 2012).

3.2.2.3 Design

The experiment employed a mixed multifactorial (2x2x3) design. The first within-participants condition was testing sequence (first and second), the second within-participants condition was odour verbalisability (high and low), and the between-participants condition was concurrent task (concurrent articulation, mental rotation, and no task).

In the mental rotation task, the congruent and incongruent trials were balanced in each 26-trial testing block, and the level of rotation and block length of the objects was evenly distributed. Presentation of each image was randomised within these testing blocks.

3.2.3 Procedure

Written consent was gained from all participants. The n-back procedure followed that described for Experiment 1, and the task took approximately 30 minutes for all participants. Prior to testing, participants performed an 8-trial picture version of the n-back task to demonstrate understanding of the procedure. In a quiet, well-ventilated room, participants then performed the olfactory 2-back task, separated from the experimenter and odorants by a wooden occlusion screen. Participants made their 2-back decision using a 7-button Cedrus Response Box, pressing the left button for a ‘No’ response and the right button for a ‘yes’ response. Responses were recorded using

Superlab 5. The key manipulation introduced into Experiment 2 was the between-participants inclusion of concurrent tasks (Figure 4).

3.2.3.1 Concurrent articulation group

Participants (*n* = 24) were required to repeatedly count (1, 2, 3, 4, 1, 2…) in the 8-second interval between n-back odour presentations. That is, participants were presented with an odour and made a 'yes' or 'no' 2-back match response. They then immediately counted out loud at a rate of approximately 2 digits per second, until presentation of the next odour.

3.2.3.2 Concurrent mental rotation task

In the 8-second interval between odour presentations, participants (*n* = 24) performed a visual mental rotation task. The comparison task was presented after an n-back decision, and disappeared once a congruency decision was made. Participants made a 'yes' or 'no' response based on whether the target object matched the baseline. The congruent and incongruent trials were balanced in each 26-trial testing block, and the level of rotation and block length of the objects was evenly distributed. Presentation of each image was randomised within these testing blocks.

3.2.3.3 Control group

The control group (*n* = 24) performed the n-back task, and were not required to perform a secondary task in the interval between odours.

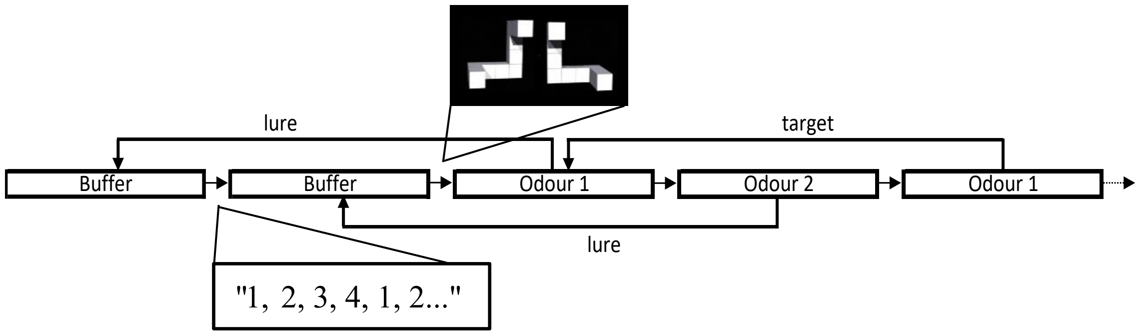

Figure 4. Schematic figure of the n-back procedure with dual-tasks. Participants were allocated to a group that, during the inter-trial interval, performed one of the following: counting task, mental rotation task, and no concurrent tasks.

3.2.4 Results

Figure 5(A-C) displays recognition sensitivity (*A'*), the proportion of hits, and the proportion of false alarms across the three concurrent task groups and collapsed across sequence number. The sequence variable was not shown in the figures because, as can be seen below, there was evidence against any main effect or interaction with sequence number. There is some variation across hits and false alarms that require addressing, however, so all three dependent variables are reported in this section.

3.2.4.1 A' sensitivity

Figure 5(A) shows the mean hit rate across the three testing groups and two odour verbalisability conditions, collapsed across sequence number. A mixed 3-factor (2x2x3) ANOVA was conducted, where the first within-participants factor was testing sequence (first and second), the second within-participants factor was odorant verbalisability (high and low), and the between-participants factor was concurrent secondary task (quiet, concurrent articulation, and concurrent rotation).

The analysis of *A'* revealed a significant main effect of verbalisability on recognition sensitivity, $F(1, 69) = 10.71$, $p = .002$, $\eta_p^2 = .13$. Performance was poorer for low verbalisability odours (*M* = 0.76, *SEM* = 0.01) than high (*M* = 0.80, *SEM* = 0.01). There was no significant change in this sensitivity between sequence 1 and sequence 2, $F(1, 69) = 0.33$, $p = .327$, $\eta_p^2 = .01$, nor did performance significantly differ across groups performing different concurrent tasks, $F(2, 69) = 0.91$, $p = .408$, $\eta_p^2 = .03$.

The predicted 2-way interaction between odorant verbalisability and concurrent task group was non-significant, $F(2, 69) = 0.26$, $p = .769$, $\eta_p^2 = .01$, as was the predicted 3-way interaction between testing sequence, odorant verbalisability, and concurrent task group, $F(1, 69) = 0.26$, $p = .771$, $\eta_p^2 < .01$. There was also no significant interaction across sequence number as a function of concurrent task, that would have indicated practice effects dependent on the secondary task performed, $F(2, 69) < 0.45$, $p = .637$, $\eta_p^2 = .01$, and the interaction of odour verbalisability effects and sequence was non-significant, $F(1, 69) = 0.26$, $p = .613$, $\eta_p^2 < .01$. Bayesian ANOVA indicated strongest support for a model with only a main effect of odour verbalisability (*BF* = 15.05 vs a null model), and that this model was strongly preferred over all interaction models.

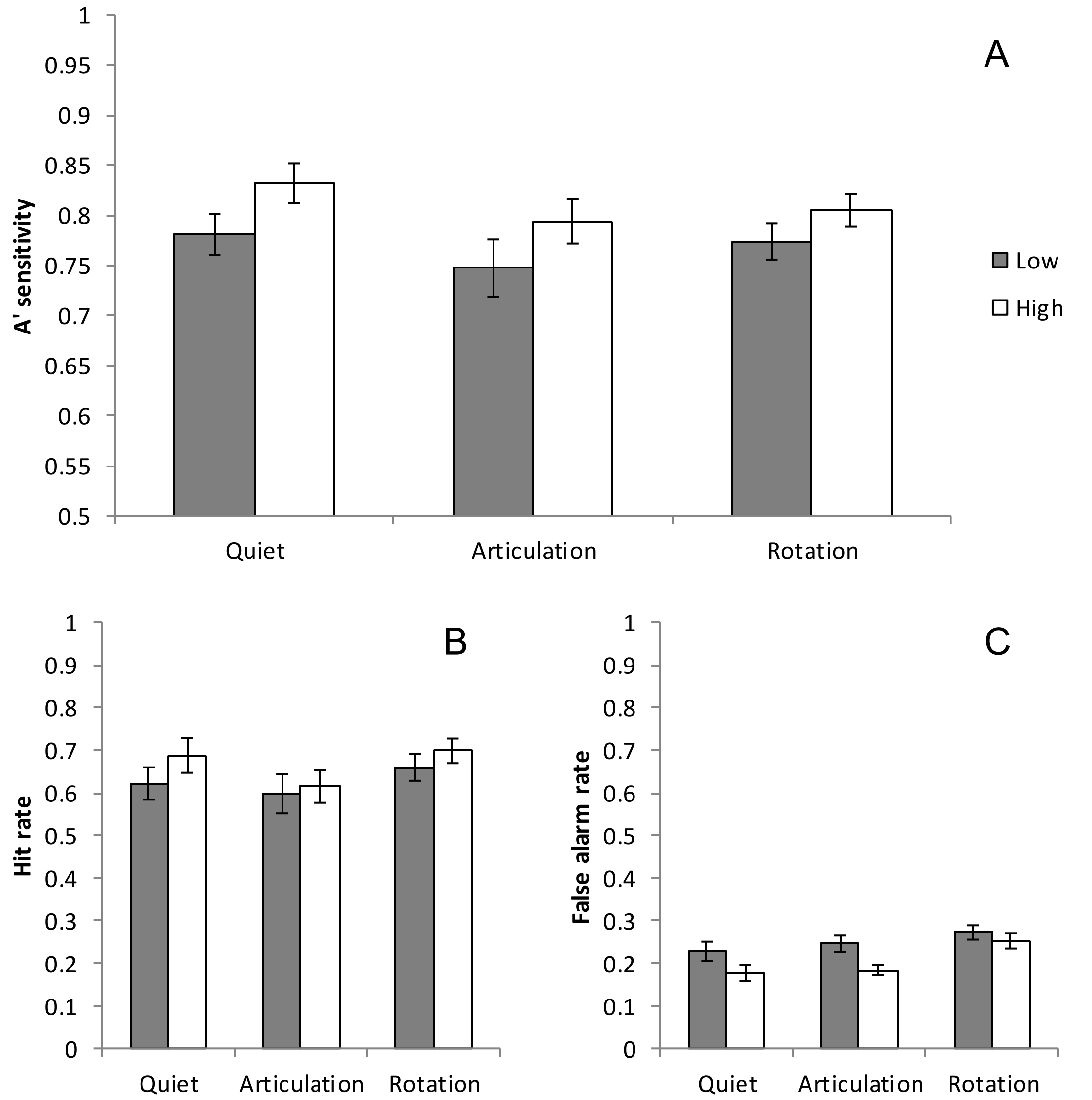

Figure 5. The mean (A) *A′* sensitivity, (B) hit rates, and (C) false alarms, for low and high odorant verbalisability, across the three concurrent task groups.

3.2.4.2 Hit rates

Figure 5(B) shows the mean hit rate across the three testing groups and two odorant verbalisability conditions, collapsed across sequence number. The ANOVA for hit rates revealed a non-significant main effect across concurrent task groups, $F(2, 69) = 1.35$, $p = .267$, $\eta_p^2 = .04$. Furthermore, there was a non-significant main effect between low ($M = .63$, $SEM = .02$) and high verbalisability ($M = .67$, $SEM = .02$) odours, $F(1, 69) =$

3.11, p = .082, η_p^2 = .04, and a non-significant main effect of sequence, $F(1, 69) = 0.53$, $p = .467$, $\eta_p^2 = .01$.

There was no interaction between sequence number and dual-task groups which might have indicated an effect of concurrent task on general exposure effects, $F(2, 69) = 0.61$, $p = .545$, $\eta_p^2 = .02$. Furthermore, there was a non- significant interaction between odour verbalisability and concurrent task group, ($F(2, 69) = 0.34$, $p = .690$, $\eta_p^2 = .01$), and in contrast with Experiment 1, the interaction between odour verbalisability and sequence number was also non-significant, $F(1, 69) = 2.11$, $p = .151$, $\eta_p^2 = .03$. Importantly, these lack of significant interaction effects were not masked by a three-way between odour verbalisability, sequence number, and concurrent task group ($F(2, 69) = 0.40$, $p = .671$, $\eta_p^2 = .01$). This suggests that the expected main effect of verbalisability, or the interaction of verbalisability across testing sequences, was not attenuated by concurrent articulation. Bayesian ANOVA revealed all models to be in favour of the null, though a verbalisability main effects model was insensitive (1.35 in favour of the null). Together, these findings support no effect of any manipulation on hit rates, aside from insensitive evidence against a main effect of verbalisability.

3.2.4.3 False alarm rates

Figure 5(C) shows the false alarm rates for low and high verbalisability odorants, across the three concurrent-task groups. Analysis revealed a significant main effect of concurrent-task group, $F(2, 46) = 3.97$, $p = .023$, $\eta_p^2 = .10$, indicating changes in the number of false alarms depending on the concurrent task that was performed. Comparisons between groups using Tukey post-hoc tests revealed significantly greater false alarms ($p = .027$) in the concurrent rotation task group ($M = .26$, $SEM = .02$) than the control group ($M = .20$, $SEM = .02$). A non-significant difference was observed ($p = .093$) between false alarms in the rotation task group and the concurrent articulation group (M = .22, SEM = .02), and a non-significant difference was found between concurrent articulation and quiet groups (p = .850). A significant main effect of odour verbalisability was also observed, where responses to low verbalisability odours saw significantly greater false alarms (M = .25, SEM = .01) than high verbalisability odours (M = .20, SEM = .01), $F(1, 69) = 18.14$, $p < .001$, $\eta_p^2 = .21$. A non-significant main effect of sequence observed, $F(1, 69) = 3.02$, $p = .087$, $\eta_p^2 = .04$, and this did not interact with concurrent task group, $F(2, 69) = 0.90$, $p = .413$, $\eta_p^2 = .03$. Unlike Experiment 1 verbalisability did not interact with sequence number, $F(1, 69) = 1.02$, $p = .316$, $\eta_p^2 = .02$, Importantly, the predicted interactions between odour verbalisability and concurrent task group, ($F(2, 69) = 1.42$, $p = .250$, $\eta_p^2 = .04$), and between odour verbalisability, sequence number, and concurrent task group ($F(2, 69) = 0.13$, $p = .882$, $\eta_p^2 < .01$) were non-significant.

Bayesian ANOVA revealed the best model to contain main effects of verbalisability and concurrent task group (1085.72 vs. the null model). This model got substantially worse by inclusion of all main effects, or the inclusion of any interaction terms. Together, these findings suggest evidence against any interaction with concurrent task group, and with sequence.

3.2.5 Discussion

Experiment 2 has further replicated the n-back advantage for verbalisable odours (see Experiment 1 and Jönsson et al., 2011). However, contrary to the prediction, CA did not attenuate the superior recognition for verbalisable odours. The proposal was that the advantage for verbalisable odours resulted from verbal labelling and rehearsal of those labels, and yet, supposed disruption of this process through CA did not reduce this verbalisability benefit. This suggests that for even these high verbalisability odours, it is not the use of a verbal rehearsal strategy that is responsible for the n-back performance.

Moreover, in contrast to Experiment 3.1, the advantage for verbalisable odours was found across sequences. Assessment of *BF* suggested the data were evidence against an interaction effect, indicating superior recall that did not emerge following refinement of verbal codes.

The null effect of CA is consistent with some previous work on olfactory short-term memory, and provides some support for a proposed olfactory-specific storage buffer (e.g. Andrade & Donaldson, 2007; cf. Annett & Leslie, 1996). That is, maintenance of odours in working memory is not dependent on verbal rehearsal, and instead might be processed as some other, perceptually-based, code. A caveat of these interpretations, however, is that they are dependent on the assumption that concurrent counting is effective in impairing verbal rehearsal processes. Whilst there exists evidence supporting the elimination of verbal rehearsal during the n-back task following concurrent articulation (Vuontela et al., 1999), it has been argued elsewhere that concurrent articulation does not have a modality-specific effect on articulatory rehearsal (e.g. Jalbert, Neath, & Surprenant, 2011). Instead, concurrent articulation may have a general disruptive effect on working memory, corrupting modality-independent features by adding noise to the memory trace.

Since CA did not remove the verbalisability advantage, one might argue that the benefit for those high verbalisable odours is derived not from verbalisability per se, but a correlate of that dimension. One candidate dimension is familiarity, for which Chapter 2 reports a strong positive correlation with verbalisability (r = .88; close alignment between familiarity and verbalisability is also reported in Jönsson et al., 2011). Moreover, Jönsson et al. found that the more nameable/familiar odours were easier to discriminate (although discriminability was not able to completely account for the variance between high and low verbalisable odours). Rather than verbal rehearsal, it is possible that the high nameable/familiarity odours are maintained using non-verbal rehearsal/refreshing processes (e.g. Raye et al., 2007). However, in *A'* scores a concurrent mental rotation task also did not selectively impair high verbalisable odour performance, which would be expected if an executive refreshing process was being applied to these odours. Indeed, Bayesian evidence against an effect of concurrent rotation was preferred by a factor of 6.65. However, this finding was not replicated across all dependent variables, with a main effect of concurrent rotation observed on the ability to reject lure items (false alarms). Consequently, though there was no effect on overall recognition sensitivity, there was some effect of a concurrent task that loaded executive resources (Simmons, 2000) that means some maintenance process cannot be ruled out. Why this concurrent rotation task only increased false alarms is unclear, but is speculated to be due to limited control processes that allow the distinction between close-lures and targets (Kane et al. 2007).

It should also be considered that although the n-back task is typically considered a task containing maintenance and manipulation (Ragland et al., 2002; Watter, Geffen, & Geffen, 2001), other proposed *low control* strategies may support retrieval of probe item and an estimate of its serial position without the need for rehearsal (Juvina & Taatgen, 2007; B McElree, 2001). These strategies suggest participants do not maintain items during the retention interval, but instead an item that is determined to be familiar is then matched to a time estimate (Juvina & Taatgen, 2007), or retrieved in a serial search where one item cues retrieval of the next (B McElree, 2001). A reliance on these strategies for olfactory information would be consistent with previous suggestions that the rehearsal of olfactory information is difficult, or impossible (Stevenson, 2009). However, the (albeit weak) evidence for an effect of concurrent rotation in the retention interval suggests that some maintenance process is being performed.

It is generally agreed that the experimental familiarity of a probe item can contribute to an *n*-back decision (Harbison et al., 2011; Juvina & Taatgen, 2007; Ralph, 2014). Consequently, failure to recollect an item may lead to a decision based on this familiarity signal, whether as part of a control strategy to establish the age the probe's presentation, or as a relatively automatic rejection based on a familiarity criterion (Juvina & Taatgen, 2007). Evidence for recognition-without-identification in odours (Cleary, 2010), and dissociated brain activations when odours can or cannot be named (Zelano et al. 2009) suggest differing strategies may be adopted for the two groups of odours in this task. For example, participants may be able to adopt a high-control working memory strategy for odours that are identified, because the availability of semantic information facilitates participant ability to generate an internal representation of an odour (Tomiczek & Stevenson, 2009). In contrast, recognition-without-identification may occur for the low verbalisability odours, and result in the use of a low-control or purely familiarity-based strategy (Juvina & Taatgen, 2007). However, the general effect of concurrent rotation provides tentative evidence for the same strategy across odours, meaning the differences may instead be based on a quantitative increase in ability to perform the task.

Alternatively, the performance advantage may be explained by the perceptual advantages familiarity to an odour provides. For example, familiar items are proposed to result in a stable and specific activation of a stored olfactory representation (Stevenson & Mahmut, 2013a; Wilson & Stevenson, 2006), which may result in discriminability improvements (e.g. Rabin, 1988), or require fewer working memory resources when performing the task (Reder, Liu, Keinath, & Popov, 2015). Experiments 3 and 4 consider the explanations offered in more detail. Specifically, Experiment 3 explores the use of recollection-based (high-control) or familiarity-based (low control) strategies across odour verbalisability, whilst Experiment 4 attempts to induce familiarity to the odours through a preliminary familiarisation session, intended to improve odour discriminability.

3.3 Experiment 3: N-back recollection and familiarity processes

3.3.1 Introduction

Experiment 3.3 applies the remember/know paradigm to the olfactory n-back task to examine the contribution of automatic and controlled retrieval processes, and how these are used to underpin n-back judgments for high and low verbalisable odours. Description of the experiment is prefixed by a summary of the purported distinction in memory between recollection and familiarity.

3.3.1.1 Recollection and familiarity in episodic memory

There is good agreement in the memory literature that processes of recollection and familiarity both contribute to a recognition judgement (Wixted & Mickes, 2010; Yonelinas, 2002). A recollection process is typically described as retrieval of qualitative information (e.g. temporal or spatial source details, or elaboration at encoding) about a presented stimulus (Gardiner, Ramponi, & Richardson-Klavehn, 1998; Jacoby, 1991; Tulving, 1985; Yonelinas, 1999). In contrast, a recognition decision based on familiarity involves evaluation of a graded strength signal, similar to that described in signal detection theory (Yonelinas, 1999, 2002). That is, familiarity-based recognition is a quantitative assessment of prior experience based on whether a signal exceeds a response criterion (Yonelinas, 1999).

The *butcher-on-the-bus* phenomenon describes the experience of seeing a highly familiar face and knowing that you recognise them, despite not remembering from exactly where (Mandler, 1980). It is this commonly reported occurrence, of high-confidence recognition in the absence of recollection, that has prompted dual-process models which assume distinct and separable processes of familiarity and recollection (see Yonelinas, 2002, for a review). Multiple procedures have been developed that estimate these recollection and familiarity processes within a task, and experimental manipulations applied to these procedures attempt to dissociate the two forms of recognition. A recollection process is considered an intentional use of memory, whilst familiarity a relatively automatic process (Jacoby, 1991). It is measures of recollection that have been shown to be affected by ageing (Koen & Yonelinas, 2016), divided attention (Dudukovic, Dubrow, & Wagner, 2009; Jacoby, 1991), and levels of processing (Gardiner, Java, & Richardson-Klavehn, 1996; Olsson et al., 2009). Furthermore, familiarity and recollection will differ in terms of processing speed (e.g. Yonelinas & Jacoby, 1994). For example, an item recognition task will be performed faster than a task that requires a judgement of the list from which an item was presented (Hintzman, Caulton, & Levitin, 1998; Yonelinas, 2002). In summary, there are dissociations shown that support a dual-process theory where familiarity and recollection are independent processes.

Dissociated effects on recollection and familiarity do not, however, always converge when using different estimation methods (Prull, Dawes, Martin 3rd, Rosenberg, & Light, 2006), and the independence of procedures that purportedly measure recollection and familiarity have come under scrutiny (e.g. Heathcote, Raymond, & Dunn, 2006; Wixted & Mickes, 2010). Single-process memory-strength interpretations, in contrast, suggest estimates of recollection and familiarity are simply measurements of the amount of evidence available (Dunn, 2008), though such interpretations struggle to explain the butcher-on-the-bus phenomenon. That is, high confidence familiarity without recollection is not predicted in a continuous single-process model because such an effect (high familiarity without recollection) supports dissociated memory processes (Wixted & Mickes, 2010). Instead, recent models have taken the distinction between recollection and familiarity, but described an eventual recognition decision as an aggregated sum-of-strength from the two distinct processes (Rotello, Macmillan, & Reeder, 2004; Wixted

& Mickes, 2010). Wixted and Mickes (2010) have proposed a continuous dual-process (CDP) model, which describes the combination of these processes that leads to a recognition decision. The important difference to dual-process models is that recollection is also placed on a strength scale, rather than an all-or-nothing threshold that overrides familiarity when it is achieved (e.g. Yonelinas, 1999). These models have important implications for the interpretation of recollection and familiarity estimation methods.

Procedures for estimating recollection and familiarity include analysis of receiver-operating characteristics (ROC), application of the remember-know response paradigm, and the process-dissociation procedure (PDP). This thesis focuses on the remember-know paradigm (Tulving, 1985). This response method utilises an introspective judgement from participants as to their recollective experience. That is, participants must indicate whether an *old* decision was based on retrieval of explicit details about the item, or was based on familiarity to the item. Considered within a dual-process recognition model, recollection supports r*emember* judgements, whilst familiarity without recollection supports *know* judgements (Evans & Wilding, 2012; Koen & Yonelinas, 2016; Yonelinas, 2002). In addition, some procedures allow an old judgement based on a *guess*, which may pick up responses that are made from inferences not directly related to memory for the item (Gardiner et al., 1998), or alternatively may be interpreted in a signal-detection model as the most lenient criterion for an *old* judgement (Green & Swets, 1974; Wixted & Mickes, 2010).

The proportion of remember-know decisions in a task can be independently manipulated by changing task demands. For example, the role of familiarity in a task (*know* judgements) can be selectively increased by increasing perceptual fluency (e.g. through masked repetition priming, Rajaram, 1993). In comparison, deeper encoding is associated with an increase in recollection (Gardiner et al., 1996; Yonelinas, Dobbins, Szymanski, Dhaliwal, & King, 1996), and a levels-of-processing effect has shown independent effects on *remember* judgements (Rajaram, 1993, 1998). In 1998, Rajaram examined whether *remember* judgements are related to the saliency or distinctiveness of the stimuli. Conceptual saliency was manipulated by adjusting the meaning of to-be-remembered homographs at encoding. That is, participants were given a short phrase that established the dominant meaning (e.g. *body part-CHEST*) or non-dominant meaning (*cabinet-CHEST*). The dominant meanings were designated *a priori* from previous research demonstrating preferential access to these meanings (Forster & Bednall, 1976). In a second experiment, the perceptual distinctiveness of stimuli was manipulated to be orthographically distinctive (e.g. *subpoena*) or orthographically common (e.g. *cookie*). Rajaram demonstrated increased recollection from both conceptual saliency and psychological distinctiveness, whilst no effect was observed on familiarity-based responding. This shows that recollection of stimuli can be selectively influenced by manipulations of conceptual fluency and perceptual distinctiveness.

The remember-know paradigm has been criticised, however, with concerns over whether judgements should be considered process-pure measures of recollection and familiarity. Proponents of a CDP interpretation of the remember-know task grades *remember* and *know* judgements along a strength scale, and this reflects the strength of evidence for an item's previous occurrence (Dunn, 2008; Hirshman & Master, 1997; Wixted & Mickes, 2010). They propose that a *remember* response is given if the strength of evidence exceeds a strict criterion, and a *know* response given if the evidence falls below this level but surpasses a less stringent criterion that determines whether an item should be judged as *new* (Dunn, 2004). Compelling evidence for this interpretation is shown in the relationship between *remember* hit and false alarm rates.

Specifically, a strong positive relationship between the two (i.e. as *remember* hits increase, *remember* false alarms also increase) suggest that these *remember* responses are made based on the assignment of a *remember* criterion rather than a qualitatively different memory (Wixted & Mickes, 2010; Wixted & Stretch, 2004).

Furthermore, evidence against independent *remember-know* responses can be observed through comparisons with alternative measures memory processes. For example, recollection can be estimated using receiver-operating characteristics (ROC), where the proportion of hits and false alarms are plotted as a curve according to a changing response criterion. The high confidence recognition associated with recollection means this type of memory can be estimated from the intercept of the curve and the asymmetry that occurs as this value increases (Yonelinas & Parks, 2007). The recollection component of an ROC curve can also be independently dissociated from a familiarity component, both by increasing reliance on recollection in a paired association task (Sauvage, Fortin, Owens, Yonelinas, & Eichenbaum, 2008), and increasing reliance on familiarity adding a response deadline to the task (Sauvage, Beer, & Eichenbaum, 2010). Importantly, when ROC curves were compared with remember-know judgements, the convergence between the two measures improved when participants were warned they may be asked to justify their remember decision (Rotello, Macmillan, Reeder, & Wong, 2005). That is, *remember* responses were subject to bias when participants knew that had to be sure about the details their recollection, suggesting the decision to make a *remember* decision may be based on a continuous underlying process.

The CDP model therefore suggests that though *remember* responses are associated with recollection and *know* responses with familiarity, these are not process-pure and will ordinarily reflect different degrees of memory strength. A *know* response is instead dependent on whether the decision is based primarily on familiarity or recollection, not that the decision is made in the absence of recollection. Consequently, a *know* response may include some recollection (Wais, Mickes, & Wixted, 2008). This can explain the presence of strong item recognition receiving a *know* response (i.e. the butcher on the bus phenomenon, which would not receive a *remember* response despite high recognition strength), by nature of the relative combination of familiarity and recollection.

Evans and Wilding (2012) assessed the dual-process and evidence-strength explanations for the remember-know procedure, and provided support for an independence explanation of response types. Magnetoencephalographic (MEG) indices were calculated from post-stimulus epochs at regions associated with recollection or familiarity (Bridson, Muthukumaraswamy, Singh, & Wilding, 2009), and compared between when participants responded *remember, know,* or *new.* Importantly, a dual-process explanation (e.g. Yonelinas, 1999) would predict greater recollection indices for *remember* responses and greater familiarity indices for *know* responses. That is, although familiarity may be present in a *remember* response, there is no lower limit for the level of familiarity required. In comparison, a *know* response must exceed a fixed criterion of familiarity which will average greater than that present for *remember* responses. This was compared to an evidence strength explanation (e.g. Wixted & Mickes, 2010) which would predict, in general, greater familiarity and recollection indices for *remember* compared to *know* responses. Their results supported the dual-process explanation, where recollection and familiarity MEG indices made independent contributions to *remember* and *know* judgements.

In summary, the convergence of remember-know findings with ROC and PDP paradigms is generally high, and judgements are dissociated in accordance with theories related to independent recollection and familiarity (e.g. Koen & Yonelinas, 2016). When applying this paradigm to recognition tasks it is appropriate to consider both dual-process and continuous dual-process interpretations, though they both generally suggest that a *remember* response is related to increased recollection, and a *know* response is related to increased familiarity in the absence of recollection.

3.3.1.2 Familiarity in Working Memory

Working memory is typically described in models as the cognitive control of active representations (Baddeley & Hitch, 1974; Cowan, 2008). However, measures of working memory performance, like long-term memory tasks, may be similarly influenced by both controlled and automatic process (e.g. Hedden & Park, 2003; Loaiza et al., 2015; Schmiedek, Li, & Lindenberger, 2009). A strong influence of automatic familiarity processes may explain why the n-back task and other measures of working memory do not always correlate, because variance due to differences in working memory ability may be limited to a small number of trials or absent entirely (Redick & Lindsey, 2013; Schmiedek et al., 2014). Furthermore, working memory tasks are most closely related to fluid intelligence when they constrain strategies making use of item familiarity (Schmiedek, Hildebrandt, et al., 2009), or when performance is based on an independent estimation of recollection (Loaiza et al., 2015).

Whilst there appear to be similar processes of recollection, familiarity, and their combination to form a memory decision, the sources of evidence in episodic memory and working memory may differ (Göthe & Oberauer, 2008). Recollection in episodic memory reflects retrieval from long-term memory, whereas recollection in working memory tasks reflects retrieval from working memory. In the n-back task, this means a controlled retrieval of the n-back item, and this information is then used to determine whether to accept or reject the item as a target. Recollection in the n-back task is therefore associated with a successful updating and maintenance process, perhaps as the consequence of successful binding between probe item and its context (Oberauer, 2005). Furthermore, familiarity in working memory refers to the residual activation of representations (e.g. Cowan, 1999), compared to familiarity in long-term memory based on the ease with which items are reactivated (Göthe & Oberauer, 2008).

In the n-back task, the role of familiarity differs further still, because after multiple trials all items will be familiar and so the decision is not simply whether an item is 'old'. Schmiedek, Li, et al. (2009) and Oberauer (2005) describes familiarity in the n-back task to denote an automatic source of evidence about the extent the probe item matches activated representations in long-term memory, and this returns a strength signal which can be accepted or rejected based on a criterion. However, the inclusion of recently-presented lures makes this familiarity-based strategy unreliable for target acceptance. Indeed, a reliance on familiarity-based responding in older participants is reflected in increased false alarms for recent-lure items (Schmiedek, Li, et al., 2009). However, although the n-back task clearly contains some automatic and controlled processes, particularly when rejecting non-recent lures, it was noted earlier that constraining n-back performance to an index of only targets and recent lures does not improve the task's relationship with other working memory measures (Kane et al., 2007).

Alternatively, models of the n-back task have been described that offer a more nuanced strategy that makes use of this activation-strength signal (Juvina & Taatgen, 2007). That is, rather than simply accepting or rejecting an item based on the familiarity-strength, a low-control strategy suggests that participants will assess the temporal distance of the probe item by evaluating the activation strength of the item, and compare it to an expected activation strength of a target. Essentially, this is a controlled assessment of a familiarity signal rather than the blanket acceptance of items that exceed a response criterion.

The estimation methods discussed in terms of episodic memory have seen success applied to several measures of working memory capacity. The work of Loaiza et al. (2015), for example, has showed simultaneous recollection and familiarity processes in complex span tasks using the process dissociation procedure (PDP) (Jacoby, 1991). The recall test in complex span was manipulated to require either the previously presented items to be reported (inclusion trials), or to report the digits that were not previously presented (exclusion trials). Because recollection is required in exclusion trials, any intrusion errors are assumed to be due to familiarity in the absence of recollection. Consequently, a recollection estimate can be calculated by subtracting exclusion trial errors from inclusion trial performance. Familiarity estimates are then calculated by dividing exclusion errors by the inverse of the recollection estimate. Using this method, Loaiza et al. demonstrated not only that automatic and controlled processes contribute to performance in a working memory recall task, but also that the estimates can be dissociated by manipulating presentation times. That is, like episodic memory (e.g. Jacoby, 1998), recollection was increased with processing time, whilst familiarity was unaffected.

Though there has been, to this researcher's knowledge, no assessment of recollection and familiarity in olfactory working memory, several experiments have investigated the two processes in human and animal olfactory long-term memory. For example, ROC curves were plotted for rats during olfactory recognition tasks by adjusting the reward offered from selecting target odours or rejecting lure odours. This revealed ROC curves that were remarkably similar to those observed in humans (Fortin, Wright, & Eichenbaum, 2004; see White et al., 2015 for a full discussion of this series of experiments). By adding an associative-memory component to the task (increasing recollection), or by manipulating the speed requirements for a response (increasing reliance on familiarity), they also demonstrated that these recollection and familiarity components of the ROC curve can be dissociated.

In human olfactory memory, an additional concern is the influence of odour familiarity and identifiability. Larsson, Öberg, & Bäckman (2006) assessed recollective experience in olfactory memory across age, incidental and intentional learning, and odour familiarity. They demonstrated a three-way interaction, where the older adults gave proportions of *remember; know,* and *guess* responses that were unaffected by odour familiarity. In contrast, younger adults gave a far greater proportion of *remember* responses compared to *know* and *guess,* to familiar odours only. Furthermore, after controlling for odour naming, these differences were removed. Consequently, Larsson et al. suggest that greater recollection of familiar odours in younger-adults is the result of age-related deficits in activating semantic knowledge (and thus, deficits in naming). In another study, Olsson et al. (2009) applied the remember-know paradigm in an olfactory episodic recognition task, and assessed dissociated effects of encoding depth and retention interval on familiarity and recollection. Recollection measured by *remember* responses was highest for words and identified odours compared to unidentified odours. Furthermore, these *remember* responses interacted with encoding depth, where *remember* responses were greater with deeper encoding for only identified odours and words. In contrast, recognition in the absence of recollection, measured by *know* responses, was stable across encoding and stimulus conditions. Consequently, they suggested that memory for identified odours more closely resembled memory for words than unidentified odours, and that mapping of an odour percept to semantic knowledge benefits ability to recollect the item. These findings may suggest that unidentified odours are qualitatively different in respect to the utilisation of familiarity and recollection. However, Olsson et al. instead propose the conceptual salience interpretation from Rajaram (1998) as an explanation for their findings. That is, identified odours are more conceptually salient than their unidentified counterpart, which directly influences the ability of participants to recollect these odours. In summary, despite several possible interpretations, these studies show that the remember-know paradigm in olfactory episodic memory ascribes an advantage to recollective experience when odours can be named.

The application of the remember-know paradigm to the n-back procedure is, to the researcher's knowledge, the first application of this metacognitive measure in this working memory task (but see Schmiedek, Li, et al., 2009 for analysis of familiarity in the n-back task using recent-lures). It should be noted that application of the remember-know paradigm deviates from its typical employment in tests of recognition memory. In those tasks (Evans & Wilding, 2012; Koen & Yonelinas, 2016; Tulving, 1985; Yonelinas, 2002), a 'yes' response is made (*old*), and a *remember* judgment (*R*) provided when the participant recollects contextual details of the previous exposure. However, in the n-back task, an *old* recognition judgement alone is insufficient to perform the task, and therefore the 'recognition' judgment becomes '*do I recognise this item as the odour presented two previous?*' In addition to identity information, an n-back recollection (*R*) judgment requires positional recall, presumably as result of a successful binding between item and context (Oberauer, 2005).

For a *know* response (*K*), participants are making a recognition judgment based upon familiarity. The typical interpretation of this response would be a decision regarding whether this strength signal exceeds a particular criterion. In the n-back task, *K* responses are expected to pick up responses made using this strategy. However, in addition, a low-control strategy has been described which compares activation strength of an item to a strength estimate of a target item. These are also expected to be picked up by *K* responses, due to the lack of contextual details retrieved and the reliance on familiarity to make this decision.

In addition, a *G* (guess) response is included in the present study for situations in which a correct response was made in the absence of any recollective experience or response strategy (Gardiner, Ramponi, & Richardson-Klavehn, 2002). However, it should be noted that such a response may instead be explained under the CDP model as a particularly weak memory (Wixted & Mickes, 2010).

The present experiment will replicate the method of Experiment 1, with the additional requirement that participants will be instructed to provide a *K (know)*, *R (remember)*, and *G (guess)* judgement following any 'yes' responses. If, as suggested above, high verbalisability odours are more amenable to a control strategy wherein the test odour is recollected related to its serial position, one might predict a greater proportion of hits that receive a *remember* response when compared to the low odour group. In contrast, one might predict a greater reliance on *know* responses for the low verbalisable odours due to the employment of a familiarity-based strategy.

3.3.2 Method

3.3.2.1 Participants

Twenty-four female Bournemouth University students (mean age = 20.21, *SD* = 3.19) participated in exchange for course credit. The same exclusion criteria as described for Experiments 1 and 2 were applied. None had participated in Experiments 1 or 2. Ethical approval was obtained from the Bournemouth University Ethics Committee.

3.3.2.2 Materials

The olfactory stimuli were taken from the same corpus of odours described for Experiment 1. Twelve odours were selected based upon normative verbalisability scores reported in Chapter 2 (see Appendix B). Six odours were high verbalisability (M = 2.64, SD = 0.11) and six were low verbalisability (M = 1.34, SD = 0.26), and these differed significantly, $t(10) = 11.54$, $p < .001$, $d = 6.66$, $BF_{10} > 100$. Furthermore, the verbalisable odours were highly familiar, with a minimum normative score of 5.60, and the low verbalisable odours were unfamiliar, with a maximum normative score of 3.61. The familiarity ratings for the high verbalisability odours (M = 5.83, SD = 0.18) were significantly higher than the low verbalisability odours (M = 3.33, SD = 0.25), $t(10) = 19.87$, $p < .001$, $d = 11.47$, $BF_{10} > 100$. Detailed assessment of odour familiarity is included as a possible alternative to verbalisability. This is premised on the findings in Experiment 2 that suggest the working memory advantage is not due to increased verbal rehearsal of the items. However, like the earlier experiments, verbalisability and familiarity scores correlate highly, so is a de-facto manipulation of both familiarity and verbalisability.

Furthermore, to address a possible confound where differences in intensity between odorant sets could explain performance differences attempts were also made to more closely balance odours on normative intensity scores. A comparison of differences in intensity scores was non-significant, $t(10) = 1.81$, $p = .101$, $d = 1.04$, $BF_{01} = 0.83$, though the data were insensitive to differences. Finally, pleasantness ratings were lower for low verbalisable odours, $t(10) = 7.71$, $p < .001$, $d = 4.45$, $BF_{10} > 100$, though the hedonic strength scores across odour sets (a measure of each pleasantness rating's deviation from a neutral midpoint) did not significantly differ, $t(10) = 0.79$, $p = .449$, $d = 0.46$, $BF_{01} = 1.76$.

3.3.2.3 Design

The design was similar to Experiment 1, but also included a metacognitive measure (*K*, *R*, and *G*) following a 'yes' response. Furthermore, the number of trials were reduced by adjusting the ratio of lures to targets, such that the twelve odorants appeared as a 'target' once (33% of trials) and twice as a 'lure' (66% of trials). The reduced number of trials served to limit olfactory fatigue effects, and the number of lures were reduced to limit a bias against responding to items as targets. Each sequence therefore totalled 36 items, and in contrast to the blocked design employed in Experiments 1 and 2, were presented as mixed blocks of low and high verbalisability odours.

The mixed-block design was chosen for the present experiment because if a switch in strategy does indeed occur across odour verbalisability (either through a conscious decision to switch, or automatically as a consequence of failed recollection), participants would likely notice the sudden increase in *know* responses that correspond to the change in odour sets, and perhaps adjust their responses accordingly.

3.3.2.4 Procedure

The procedure was as described for Experiment 1, but following a 'yes' response participants were required to provide an additional metacognitive decision. Instructions for this response were a modification of that described by Rajaram (1993): an *R* response was required when participants explicitly recollected the odour and its occurrence in the correct n-back position; a *K* response was required when the 'yes' response was based on the level of familiarity associated to the item; and a *G* response was required when participants made a 'yes' decision that was based on some other reasoning, strategy, or if they were unsure why they had responded 'yes'. Responses were made on a Cedrus Response Box and the input recorded using Superlab 4.5.

3.3.3 Results

3.3.3.1 Working Memory Performance

Figure 6(A-C) displays recognition sensitivity (*A'*), the proportion of hits, and the proportion of false alarms across the odour verbalisability groups. For these analyses, hit and false alarm *guess* responses were removed (see Olsson et al., 2009 for a similar application of this method; although it should be noted that including guess responses did not change the outcome from the analysis detailed below).

Comparisons of *A'* sensitivity between low (M = .79, SD = .16) and high (M = .88, SD = .08) verbalisability odours supported improved n-back performance for the high odour set, $t(23) = 3.03$, $p = .006$, $d = .69$, $BF_{10} = 15.02$. A comparison of hit rates for low (M = .58, SD = .23) and high (M = .74, SD = .16) verbalisability conditions also revealed significantly greater hits for verbalisable odours, $t(23) = 3.71$, $p = .001$, $d = 0.85$, BF_{10} = 61.77. However, false alarm data (the proportion of incorrect responses to lures) revealed a non-significant difference, $t(23) = 1.01$, $p = .323$, $d = 0.19$, $BF_{10} = 0.56$, with anecdotal support for the null hypothesis.

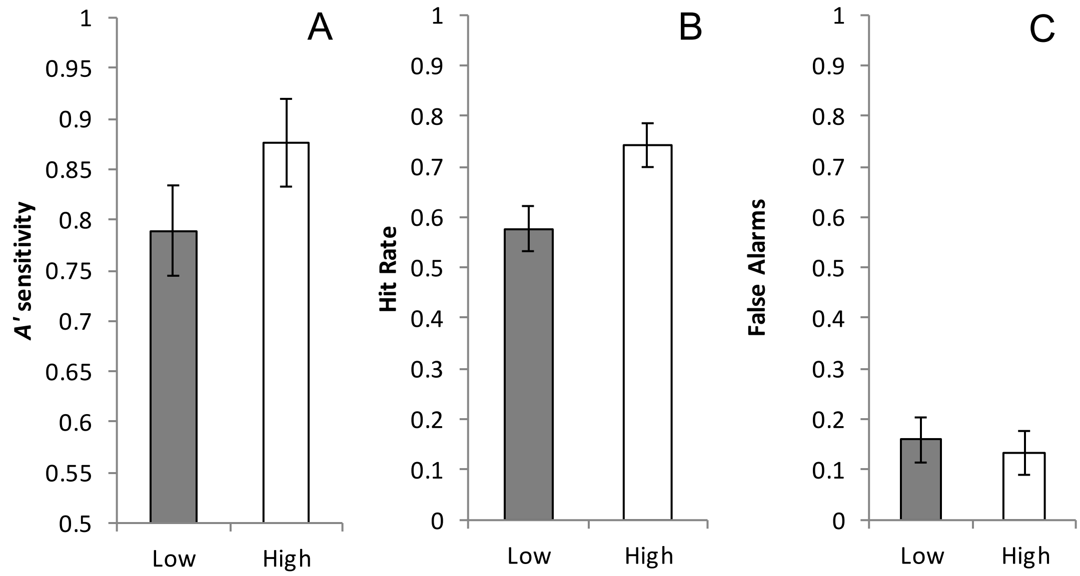

Figure 6. Low and high verbalisability odour (A) *A′* sensitivity, (B) hit rates, and (C) false alarm rates. Error bars represent standard error of the mean.

3.3.3.2 Metacognitive Responses

The proportion of *remember, know,* and *guess* responses were calculated from the number of 'yes' responses given by each participant, separately for targets and lures. This is a *relative* calculation that gives the proportion of a response type without consideration to the number of responses given (see Larsson et al., 2006, for an example of this analysis applied to remember-know responses). An alternative analysis of recollective experience using *absolute* proportions based on the total number of trials within a condition revealed the same pattern of results (applied in Olsson et al., 2009), and are not reported.

3.3.3.3 Hits

The proportion of *remember, know,* and *guess* responses were calculated from the number of correct 'yes' responses in order to examine whether the proportion of response types differed as a function of odour verbalisability. Figure 7(A) shows the proportion of response types for these correct target responses. A 2-factor (2x3) repeated-measures ANOVA was conducted where the first manipulation was odour verbalisability (low and high) and the second manipulation was metacognitive judgment (*K, R,* and *G*). The main effect of odour verbalisability was not assessed because the sum of $P(Hit_K) + P(Hit_R) + P(Hit_G) = 1$ for both low and high verbalisability sets. The main effect of response type was significant, $F(2, 46) = 3.19$, $p = .050$, $\eta_p^2 = .12$, as was the theoretically important interaction between odour verbalisability and metacognitive judgment, $F(2, 46) = 6.51$, $p = .003$, $\eta_p^2 = .22$. A Bayesian ANOVA indicated strong support for a model that included a response-type main effect and an interaction between verbalisability and response type (*BF* = 389.56 vs a null model). This model was preferred to a response type main effect model by a factor of 90.01. That is, there was strong evidence for an interaction between odour verbalisability and the type of response.

In order to examine this interaction in more depth, differences between the proportion of responses for low and high verbalisability odours were compared independently for each response type. Paired t-tests supported a hypothesis of lower *G* responses for high verbalisability odours ($M = .19$, $SD = .20$) compared to low verbalisability odours ($M = .39$, $SD = .22$), $t(23) = 5.19$, $p < .001$, $d = 0.96$, $BF_{10} > 1{,}000$. In contrast, there was evidence against greater *K* responses for low verbalisability odours, $t(23) = -0.40$, $p = .695$, $d = 0.11$, $BF_{10} = 0.16$. Finally, a hypothesis of greater *R* responses for high verbalisability odours ($M = .51$, $SD = .24$) compared to low verbalisability odours ($M = .33$, $SD = .25$) was supported, $t(23) = 2.50$, $p = .020$, $d = 0.71$, $BF_{10} = 5.42$.

3.3.3.4 False alarms

The proportion of false alarms for each response type were analysed across odour verbalisability in a separate (2x3) ANOVA, and is shown in Figure 7(B). This revealed a main effect of response type, $F(2, 46) = 3.69$, $p = .033$, $\eta_p^2 = .138$, which did not interact with odour verbalisability, $F(2, 46) = 0.56$, $p = .577$, $\eta_p^2 = .024$. The preferred model contained only a main effect of response type ($BF = 18.14$ vs. a null model), and was preferred over a model with an interaction by a *BF* of 21. However, Bonferroni-adjusted pairwise comparisons collapsed across verbalisability conditions revealed non-significant differences between *K, R*, and *G* false alarms. Analysis with Bayes Factors revealed evidence against a difference between *K* and *R* responses ($BF_{01} = 4.14$), but did reveal some evidence for a difference between *G* and *R* responses ($BF_{10} = 4.00$), and anecdotal support for a differences between *G* and K responses.

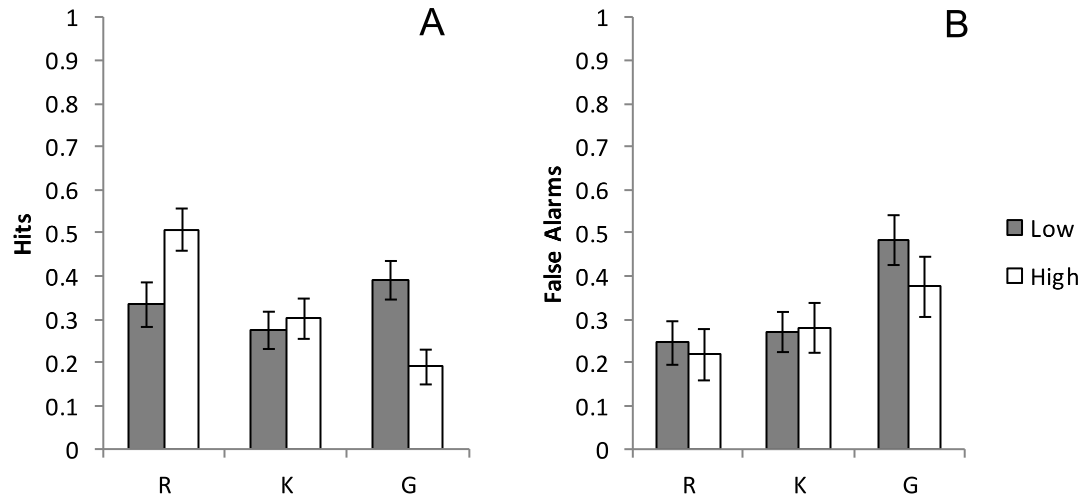

Figure 7. Proportion of metacognitive response types for (A) hits and (B) false alarms across odour verbalisability.

3.3.4 Discussion

Experiment 3 has replicated the previous finding that n-back performance is superior for high verbalisability odours, and also replicates this verbalisability advantage in a mixed-trial design (as used in Experiment 1 of Jönsson et al., 2011). The inclusion of *know/remember/guess* judgments revealed an interaction that suggests different contributions to a response across odour verbalisability. The proportion of *remember* hits reflects correct responses that were made with recollection of contextual details from study, and was higher for verbalisable odours than low verbalisability odours. In contrast, there was no difference between the proportions of *know* responses between odour sets. That is, after making a correct 'yes' response, participants deemed a similar proportion of responses from both odour sets to have been based on item familiarity in the absence of recollection.

As discussed in the Introduction to this experiment, recollection is indicative of controlled information processing in working memory (e.g. Loaiza et al., 2015). Consequently, a *remember* response not only reflects recognition of the presented item as old, but also suggests a controlled strategy that enables clear recollection of the item's assignment to the n-back position (e.g. Chatham et al., 2011; Juvina & Taatgen, 2007; Oberauer, 2005). The aim of a *know* response was to catch correct responses that were made based on familiarity in the absence of recollection. Whilst decisions based on a familiarity criterion can produce a relatively automatic contribution to recognition, prior models of n-back performance suggest that such a signal may also be assessed using a low-control time-tag strategy (e.g. Juvina & Taatgen, 2007). It should be noted that *know* responses do not distinguish the use of these strategies, as both make a judgement based on familiarity-signal. However, a reliance on a familiarity-based strategy would have been evidenced by a larger proportion of *know* hits and false alarms, which was not observed. Consequently, whilst these findings do not demonstrate the predicted shift towards a familiarity based strategy for low verbalisability odours, they do indicate a reduced reliance on recollection (Gardiner et al., 2002) that is consistent with olfactory long-term memory findings (Larsson et al., 2006; Olsson et al., 2009).

In terms of n-back strategies, what is also evident is that the proportionally greater recollection for high verbalisable odours is shifted to proportionally greater guess responses for low verbalisability odours. This suggests a general reduction in response certainty even when responding correctly to low verbalisability odours (indeed, the proportion of guessed hits for low odours are double that of high odours suggesting a weaker memory trace, see Wixted & Mickes, 2010). Olsson et al. (2009) suggest that mapping an odour percept to semantic knowledge is advantageous for episodic recognition, evidenced by improvements in both memory and recollective experience. They propose that identification is responsible for the increase in recollection because these items are more conceptually salient (Rajaram, 1998). The present findings support a similar conclusion for working memory, where verbalisability of odours is linked to an increase in ability to recollect the odours. However, Olsson et al. also suggest that identified odours produce a verbal code that provides an additional cue to memory. Whilst it is unclear whether a familiar odour in the present experiment also includes the use of a verbal label as an additional retrieval cue, in Experiment 2 it was shown that such a cue was not rehearsed throughout the retention interval.

The present study therefore shows that participants do not abandon a control strategy for low-verbalisability odours, but instead continue with the strategy in a less successful manner. However, it should be noted that such a conclusion is dependent on the ability of *know* responses to accurately reflect the use of a familiarity-based strategy. For example, it is possible that the task requirements were confusing for participants as it requires self-awareness of the information that was used to make a decision, and this decision is more complex than for a simple old/new distinction. Indeed, it is possible that the *guess* hits were greater for low verbalisability odours due to a high proportion of responses based on a strategy that participants were not able to categorise. A further possible criticism of the present study is the use of mixed verbalisability lists. Although the use of mixed-blocks was necessary to prevent participants realising from their responses that a shift in odours had occurred, it is possible that the use of this design resulted in the perseverance of a strategy throughout the task. In comparison, the blocked design in experiments 1 and 2 may have facilitated switching due to clearly defined periods of verbalisable/non-verbalisable odours. Notwithstanding these potential issues, the findings have supported previous links between recollection and odour verbalisability in this area (e.g. Larsson et al., 2006; Olsson et al. 2009).

Manipulations of task demands that affect recollection are typically associated with attentional resources that affect the encoding of stimuli (e.g. Jacoby, 1998). In an n-back working memory task however, Oberauer (2005) describes recollection as a reflection of the successful coordination of a binding/unbinding process (updating). A failure in this updating process may explain the lower recollection and subsequent increase in guessing when odours were difficult to verbalise. Indeed, the properties of high verbalisability odours include increased conceptual saliency (Olsson et al., 2009), which are proposed to support recollection by facilitating the binding process between odour item and its context at encoding (Oberauer, 2005).

However, another influence on recollection described in Rajaram (1998) is the role of perceptual discriminability. Although Olsson et al. (2009) suggest this is similar between low and high familiarity odours, it is hypothesised here that odour familiarity can increase discriminability through perceptual learning (See Section 3.0.4), and this can also lead to a working memory advantage. That is, perceptual learning may increase perceptual distinctiveness of odours in addition to conceptual saliency, proposed through perceptual learning (Goldstone, 1988) or the ability of the olfactory system to match an input pattern to a stored representation (Rajaram, 1998; Wilson & Stevenson, 2006).

In summary, Experiment 3 provides some evidence for a quantitative change in a participants' ability to recollect odours as a function of whether odours are verbalisable, rather than a qualitative shift to a familiarity-based strategy. Specifically, recollection decreased for odours difficult to verbalise. Furthermore, it is suggested that the decline in recollection is due to issues in the maintenance and updating of bindings, due to low conceptual saliency or perceptual distinctiveness (Rajaram, 1998; Wilson & Stevenson, 2006). However, as there is little evidence for verbal rehearsal in the olfactory n-back task (see null effects of CA in Experiment 2), it is possible that the saliency/distinctiveness of the high verbalisability odours is not due to the addition of labelling per se but a covariate of verbalisability (i.e. familiarity). Experiment 4 examines the role this odour familiarity, to assess whether perceptual discriminability is responsible for the n-back advantages observed.

3.4 Experiment 4: Perceptual discriminability in working memory

3.4.1 Introduction

In Experiment 2, a working memory advantage for high verbalisability odours was demonstrated that does not appear to be underpinned by an articulatory rehearsal process. Indeed, the n-back task may have been performed without rehearsal at all, and instead performed with some low-control familiarity-based strategy (see Juvina & Taatgen, 2007). In Experiment 3, it is suggested that the performance advantage is due to a quantitative increase in ability to recollect verbalisable odours and not a qualitative change in n-back strategy. Consequently, in this experiment the proposed advantage afforded to verbalisable items in working memory are investigated through the effects of perceptual familiarity on item discriminability.

As discussed in Chapter 1, models of odour perception place particular importance on the ability to match a glomerular activation input to a stored object pattern (e.g. Wilson

& Stevenson, 2003a). That is, smelling a familiar odour is suggested to activate a representation within an olfactory-specific object store that is stable, and less redolent of other odours, compared to novel or unfamiliar odours that will not activate an exact match (Li et al., 2006; Mingo & Stevenson, 2007; Stevenson & Mahmut, 2013a; Wilson & Stevenson, 2003a). This perceptual learning is independent to an advantage gained from labelling or semantic categorisation, and will improve recognition ability and discriminability from mere exposure to a stimulus (Jehl et al., 1995, cf. 1997; Rabin, 1988; Stevenson & Wilson, 2007; Wilson & Stevenson, 2003b). A similar process is proposed for expertise effects on wine discriminability, where advantages are observed independently to learned semantic associations (Melcher & Schooler, 1996).

It should, however, be noted that there exists contrasting evidence regarding the facilitative and detrimental effects of familiarity in memory performance for other (non-olfactory) modality stimuli. Some research suggests that familiarity can exert negative effects on memory. For example, whilst recall and processing tasks do see an advantage from familiarity (as determined by frequency, see Diana & Reder, 2006), recognition memory typically shows the opposite effect where low familiarity of a word exhibits better performance (Yonelinas, 2002). The recognition advantage for low frequency words is produced specifically from greater hits and reduced false alarms compared to high frequency words (Gorman, 1961; Reder et al., 2000), and explained by confusion over the source of familiarity and recollection (e.g. Reder et al., 2000). That is, participants have problems differentiating experimental familiarity and normative familiarity, resulting in more *old* responses and thus more false alarms. In contrast, hit rates are reduced because a greater number of contextual links to the stimuli result in difficulty making a clear item recollection (Guttentag & Carroll, 1994).

Other studies have reported a facilitative effect of familiarity on memory, however, and for high frequency words this has been explained by fewer processing resources required at encoding (Diana & Reder, 2006; Reder et al., 2015). Diana and Reder postulated that before any retrieval advantage can occur for low frequency words (i.e. the lack of contextual confusions, or effect of only within-experiment familiarity), there must be a successful binding between the word and its context (see also Oberauer, 2009). Consequently if the task is sufficiently taxing there will be the encoding advantage afforded for familiar words, but no retrieval advantage for the unfamiliar words because of the lack of this content-context binding.

The rationale for the experiment is therefore premised on the proposal that familiar odours will require less working memory resources in the encoding and updating of item-context bindings in the n-back task (Oberauer, 2005; Reder et al., 2015). In their study, Reder et al. examined the proposal that less familiar items consume more working memory resources by experimentally controlling familiarity to previously unknown stimuli. Over 4 weeks, the authors manipulated the frequency with which Chinese characters appeared in visual search and paired association training sessions. N-back performance was demonstrably greater for (experimentally-induced) high frequency characters over low, which they suggest supports their assertion that encoding familiar stimuli requires less working memory resources.

In this experiment, whether odour familiarity is underpinning the performance advantage for the high verbalisable odours is investigated by experimentally increasing odour familiarity through a series of preliminary rating and discrimination tasks (see Jehl et al., 1995, 1997; Nguyen et al., 2012; Rabin, 1988; Reder et al., 2015). To be clear, Experiment 4 tests whether increased familiarity for normative unfamiliar odours can improve memory performance on the n-back task. It is predicted that prior exposure will increase n-back performance for these previously unfamiliar odours due to perceptual learning (Wilson & Stevenson, 2003a), and the subsequent working memory advantage afforded to familiar items (Reder et al., 2015).

3.4.2 Method

3.4.2.1 Participants

Forty-eight Bournemouth University students (34 females, 14 males, mean age = 20.48, *SD* = 2.23), who had not previously participated in an olfactory n-back task, participated in the present study. The same exclusion criteria from earlier experiments were applied. Ethical approval was gained through the Bournemouth University ethical procedures.

3.4.2.2 Materials

Olfactory Stimuli. Although the primary aim of this experiment was to assess the effects of familiarisation on previously unfamiliar odours, the distinction between odour sets remained categorised on verbalisability. This was to prevent confusion with the between-participants manipulation (familiarisation), and to retain a clear comparison with earlier experiments in this chapter. However, low verbalisability sets should be considered as also low familiarity, and the high verbalisability set considered as high familiarity (see full comparisons below).

The olfactory stimuli were taken from the corpus of odours described in Chapter 2. Eighteen odorants were selected for the task, with twelve odours (separated into two sets, see Appendix B) specified as low verbalisability. The remaining six odours were classed as high verbalisability from these data. The strength of evidence for differences across verbalisability and familiarity, and evidence against intensity, pleasantness, and hedonic strength differences, were calculated using Bayes Factors. There was strong evidence for a difference in verbalisability scores between the two low verbalisability sets (M = 1.44, SD = 0.47) (M =2.60, SD = 0.11) and the high verbalisability odour set, $t(16) = 5.92$, $p < .001$, $d = 2.96$, $BF_{10} = 682.41$, and also for a difference in normative familiarity between low (M = 3.26, SD = 0.24) and high (M = 5.77, SD = 0.11) sets, $t(16) = 24.17$, $p < .001$, $d = 12.08$, $BF_{10} > 1{,}000$.

Although there was evidence against no difference (i.e. evidence for a difference) between pleasantness in the two sets, $t(16) = 9.04$, $p < .001$, $d = 4.52$, $BF_{01} < 0.01$, the data provided anecdotal evidence against a difference in intensity, $t(16) = 1.10$, $p = .286$, $d = .55$, $BF_{01} = 1.56$, and hedonic strength, $t(16) = 0.46$, $p = .649$, $d = .23$, $BF_{01} = 2.18$.

Familiarisation Tasks. Using a similar familiarisation method to that used for olfactory stimuli in Sinding et al. (2015), the Self-Assessment Manikin (SAM) was applied to determine an individual's affective reaction to stimuli (Bradley & Lang, 1994). The 9-point pictorial rating scales seen in Figure 8 records how happy/unhappy (pleasure), excited/calm (arousal), and controlled/in-control (dominance) the stimulus makes the participant feel.

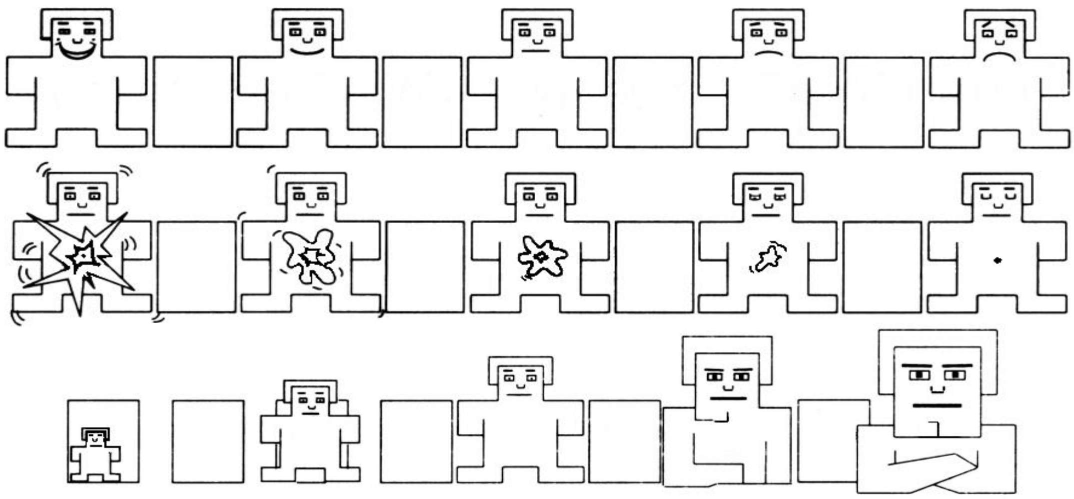

Figure 8. The Self-Assessment Manikin (Bradley & Lang, 1994)

In addition, 7-point rating scales of familiarity (highly unfamiliar/highly familiar), intensity (no odour/very intense), notes (few odour notes/many odour notes), and complexity (very simple/very complex) were used. Questions were displayed on a 22 inch monitor using open source stimulus presentation software OpenSesame (Mathôt et al., 2012) and responses were made on the keyboard number pad.

N-back task. Responses in the n-back task were recorded using Superlab 5.0 and a 7-key Cedrus RB-730 response box. The leftmost key was used for a 'no' response, and the rightmost a 'yes' response.

3.4.2.3 Design

Experiment 3 employed a mixed multifactorial (2x2) design. The between-subjects factor concerned whether participants were familiarised to odorants for use in the n-back task, during a preliminary session. The within-participants factor concerned the verbalisability of the test odours during the n-back task (low or high).

Participants were randomly allocated to the familiarised or control group at the beginning of the preliminary session. In this session, all participants were exposed to six odorants with low normative verbalisability scores (Chapter 2). The familiarised group experienced these same six odorants as part of the low verbalisability condition during the subsequent n-back task (the testing session), whilst the control group experienced an alternative odour set. In addition, all participants experienced a high verbalisability odour set which was included in the n-back testing session only. The presentation of the two low verbalisability odour sets was counterbalanced such that half of the familiarised participants experienced *low verbalisability set 1* twice, and half experienced set *low verbalisability set 2* twice. Similarly, half of the control group experienced *low* *verbalisability set 1* in the preliminary session and *low verbalisability set 2* in the testing session, and the other half experienced the opposite.

The n-back task followed the same design in Experiment 3, with two testing sequences of 36 trials, but did not require an additional metacognitive judgement following a target/lure decision. A testing sequence therefore consisted of 18 low verbalisability trials and 18 high verbalisability trials, with low and high verbalisability odours presented together. That is, low and high verbalisability odour sets were mixed. This mixed design was applied primarily to prevent participants in the familiarised group realising that the sequence of 18 trials matched the odour set presented the day previous.

The number of correct target identifications (Hits), and incorrect identifications of a lure as a target (False Alarms, FA) were analysed. In addition, an index of performance was analysed using *A'*.

3.4.2.4 Procedure

The experiment was conducted in a quiet, well-ventilated lab room at Bournemouth University. Participants were tested individually across two sessions, which were separated by a minimum of 20 hours and a maximum of 28 hours. That is, participants were tested in the day following the preliminary session, at approximately the same time as that first session. Written consent was gained before the start of the first session, and further verbal consent gained prior to the second session.

Preliminary (Familiarisation) Session. The familiarisation sessions consisted of eight olfactory rating tasks, designed to give purpose to the process whilst keeping participants naïve to the real aim of the experiment. Participants were equipped with odourless vinyl gloves and sat in front of a tray of six test tubes appropriate to their group allocation. Instructions were presented on a monitor placed behind the odorants, and directed participants to smell each odour each time they were directed to do so, and to make their judgements by pressing the corresponding number to their choice. In addition, participants were given paper instructions which included detailed guidance on the SAM rating scales, modified from Bradley and Lang (1994) (Appendix C). Importantly, participants were not required to remember the odorants they were evaluating, nor were they given any indication that the odours would be used in the future session.

Participants made responses to a single question for each odorant, before repeating the process for a new question. The question order was randomised for each participant, and the odour order randomised within questions. In order to minimise adaptation, a 30-second break was built into the program after each question.

Following the odorant rating procedure, participants performed a discrimination task for pairs of odours from within their allocated odour set. Participants sat opposite the experimenter, separated by a wooden screen, and were instructed to make a verbal 'same' decision for congruent pairs or a 'different' decision for incongruent pairs. These odour pairs were held for 2-seconds under the nose of the participant, with a 2-second inter-stimulus interval (ISI). Every combination of odours was presented in a random order, meaning each of the 6 odours was presented as a congruent pair once, and within an incongruent trial 5 times. The total number of incongruent pairs therefore totalled 15 trials, and 21 comparisons. Total testing time, including breaks, was approximately 20 minutes.

Testing (n-back) Session. The n-back procedure followed that for Experiment 3, but did not require a *remember-know* judgement. Participants were given a 5-minute break between sequences, and were encouraged to get some fresh air and have a drink of water. Total testing time (including breaks) took approximately half an hour.

3.4.3 Results

3.4.3.1 A' sensitivity

Figure 9(A) shows the *A'* scores across familiarisation group and odour familiarity. A mixed 2-factor ANOVA was conducted (2x2), where the within subjects factor was odour verbalisability (low and high), and the between subjects factor was familiarisation group (familiarised and control groups). A significant main effect of odorant verbalisability was found, $F(1, 46) = 13.84$, $p = .001$, $\eta_p^2 = .21$, where high verbalisability odorants ($M = .89$, $SEM = .01$) saw greater recognition sensitivity than the low verbalisability odorants ($M = .84$, $SEM = .01$). The main effect of familiarisation group was non-significant, $F(1, 46) = 3.37$, $p = .073$, $\eta_p^2 = .07$, and importantly the interaction between odour verbalisability and familiarisation was non-significant, $F(1, 46) = 1.60$, $p = .212$, $\eta_p^2 = .03$. The preferred model was one with both main effects (55.94 vs a null model), though this was preferred over a model with just a main effect of odour verbalisability by a factor of 1.20. Therefore, these data support a model with a main effect of verbalisability which is only anecdotally improved by the inclusion of a main effect of familiarisation.

In summary, the data provides some evidence against the hypothesis that familiarisation to odorants will improve recognition sensitivity for those same odorants in a working memory task. Indeed, in contrast to the original prediction, there was a trend towards poorer sensitivity when the task included odours from the familiarisation process.

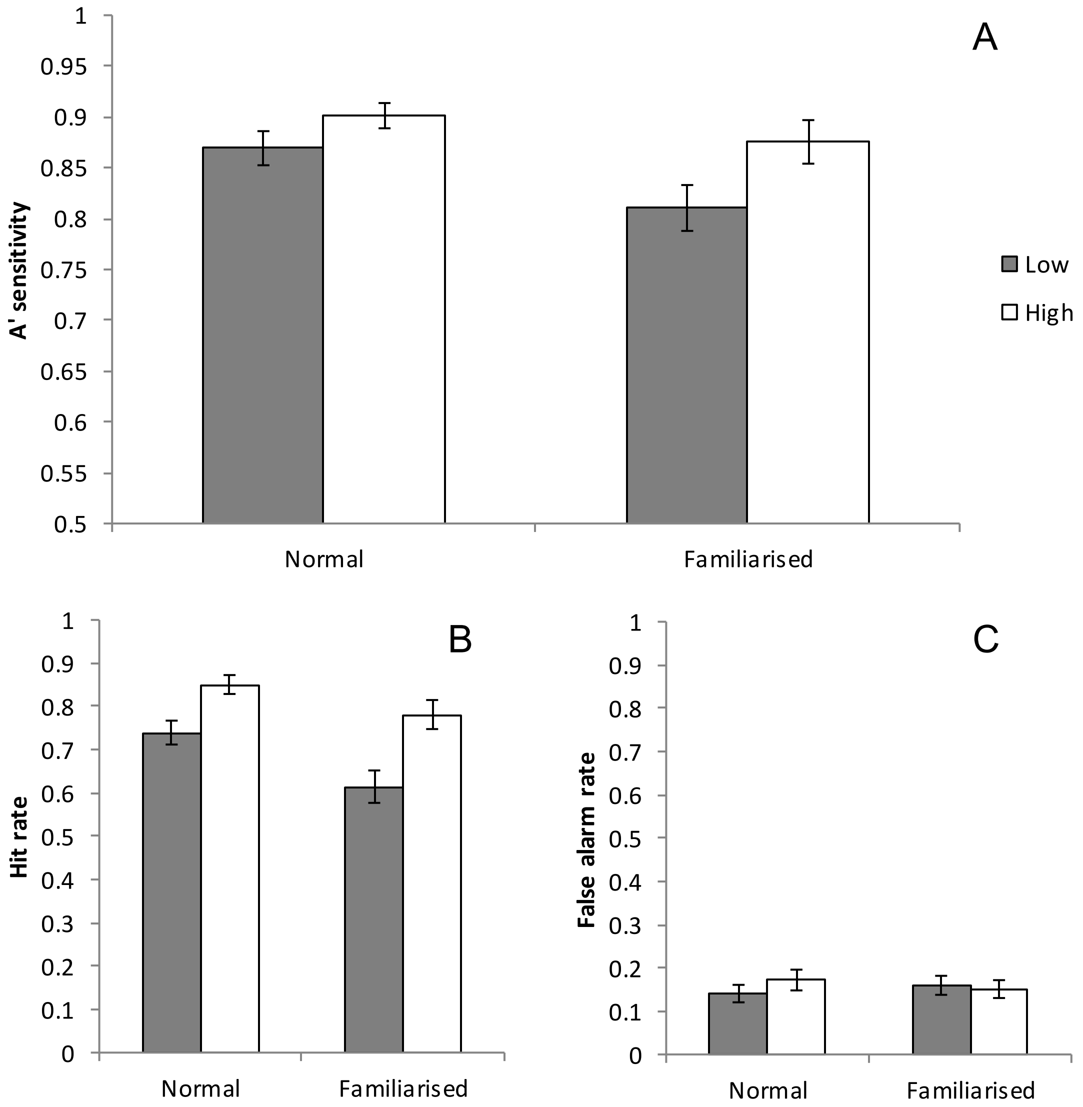

Figure 9. Working memory performance for low and high verbalisability odours as measured by (A) *A′* sensitivity, (B) hit rates, and (C) false alarms, across familiarisation groups.

3.4.3.2 Hit rates

The proportion of correct target recognition is shown in Figure 9(B), and was analysed with the same 2 x 2 mixed ANOVA. A main effect of odour verbalisability, $F(1, 46) = 36.65$, $p < .001$, $\eta_p^2 = .44$, showed greater hits for low verbalisability odours ($M = .68$, $SEM = .02$) than high verbalisability odours ($M = .82$, $SEM = .02$). In addition, a main effect of familiarisation group was also significant, $F(1, 46) = 6.95$, $p = .011$, $\eta_p^2 = .13$, and revealed *lower* overall hits for the familiarised group (M = .70, SEM = .03) over the non-familiarised group (M = .80, SEM = .03). Importantly, however, the interaction between the main effects was non-significant, $F(1, 46) = 1.47$, $p = .232$, $\eta_p^2 = .03$. A model containing both main effects was preferred (295,228.34 vs. the null model), which improved a verbalisability-only model by a factor 4.53, and was not improved by the addition of an interaction. Again, this provides evidence against an improvement in target recognition for odours that were familiarised, and instead supports a general decline in performance across the task when the low verbalisability odours had been presented in both sessions.

3.4.3.3 False alarm rate

Figure 9(C) displays the false alarm rates across familiarisation groups and verbalisability conditions. There was a non-significant main effect of odour verbalisability for false alarms, $F(1, 46) = 0.36$, $p = .552$, $\eta_p^2 = .01$, and a non-significant main effect of familiarisation group, $F(1, 46) < 0.01$, $p = .972$, $\eta_p^2 < .01$. Furthermore, there was a non-significant interaction between odour verbalisability and familiarisation group, $F(1, 46) = 1.13$, $p = .294$, $\eta_p^2 = .024$. The null model was preferred to all other models by $BF > 3$, suggesting evidence against both main effects and an interaction.

3.4.4 Discussion

Experiment 4 familiarised participants to odours which were low on normative familiarity (categorised as low verbalisability) prior to the main memory study. Whilst the main effect of verbalisability was replicated (Experiments 1-3), this experiment showed no selective improvement in working memory performance for familiarised odours in the n-back task. To be clear, an interaction was predicted such that working memory performance would improve following familiarisation but only for the low verbalisability odours with which participants were pre-exposed. This was not found. This contrasts previous findings that suggest the familiarity of a stimulus is linked to an encoding advantage (Diana & Reder, 2006), and the consumption of fewer working memory resources (Reder et al., 2015). Paradoxically, there was instead a general decline in working memory performance across the task, particularly in participant ability to accept a target item. That is, participants had more difficulty saying 'yes' to an odour that was a target if the task included odours from the prior session, and this difficulty was not localised to responses for only familiarised odorants. This unexpected pattern of results suggests some shift in responding due to the presence of familiarised items. One might suggest that the present results occurred due to two reasons; a disruptive effect of item familiarity, combined with the use of a mixed-block design for the two odour sets. However, to pre-empt the discussion below, any attempt to explain the disruptive effect of familiarisation is confounded by the superior n-back performance for the high verbalisable odours.

First, why participants familiarised to the low verbalisability odours showed a decreased hit rate is considered. In the introduction to this experiment, reduced hit rates for familiar words were considered as a result of contextual confusions that affect retrieval of the items (Reder et al., 2000). Applied to the present findings, the familiarisation of odours may have increased the likelihood of contextual confusions when comparing probe odours to the n-back item. Such an effect may have occurred independently of any encoding/updating advantage for the high verbalisability odours (Oberauer, 2005; Rajaram, 1998; Reder et al. 2015), which can explain the consistent n-back advantage for these odours seen in both groups in the present experiment, and in Experiments 1-3. However, whilst the above explanation can accommodate a decline in hits for odours that were subject to the preliminary familiarisation session, it is noteworthy that the disruptive effect of familiarisation was demonstrated as a main effect rather than an interaction. To be clear, one might predict that the effect should be detrimental to the specific odours to which familiarisation occurred; however, this was not the case as both the low and high verbalisable odours were affected.

It is possible that increased conflict between familiarity (item activation) and recollection (contextual memory of the item's appearance) resulted in some change in response strategy (e.g. Harbison et al., 2011; Juvina & Taatgen, 2007). That is, familiarised odours produce a familiarity signal that means lures cannot be rejected by familiarity alone, and a recollection process must therefore be applied to all odours to establish whether to respond to the item as a target. As a result, a conflict between familiarity and recollection that would normally occur from only recently-presented lures may have occurred for all familiarised lure items. Since targets are characterised by their recent presentation and thus strong experimental familiarity (Harbison et al., 2011), one might argue that the inflated familiarity signal disrupted the acceptance of a target by being an unreliable indicator of an item's recency. The general effect of familiarisation could therefore be attributed to the use of a mixed-block design. Specifically, unreliability of a familiarity signal is likely to have affected the response process adopted for all odours, regardless of whether the particular odour being tested was familiar due to task or pre-task familiarity. Since familiarity ceased to be an effective strategy, this may have created general confusion when responding to all odours in the mixed design.

A shift in criterion was assessed as a consequence of the disruptive effect of familiarity, and the subsequent drop in hits (see Ralph, 2014 for a discussion of strategy adoption with changing task demands). To test this, bias measure B'' was calculated from hits and false alarm rates (Stanislaw & Todorov, 1999). This is a non-parametric assessment of participant bias to respond either ‘yes’ or ‘no’ to items; a score of -1 reflects an extreme propensity to respond ‘yes’, a score of 1 reflects extreme bias to say ‘no’, and a score of 0 signifies no bias). This score was compared across groups and conditions using a 2 (familiarised or control group) x 2 (low or high odour verbalisability) ANOVA. This revealed a main effect of odour verbalisability where a stricter criterion was applied for low verbalisability odours compared to high verbalisability odours, $F(1, 46) = 13.84$, $p < .001$, $\eta_p^2 = 0.23$, but no significant main effect of familiarisation, $F(1, 46) = 1.86$, $p = .180$, $\eta_p^2 = 0.04$, and no significant interaction between this main effect and familiarisation group, $F(1, 46) = 1.90$, $p = .174$, $\eta_p^2 = .04$ (a main effect model of verbalisability only was preferred vs. a null model by a *BF* of 274.62, and did not improve with both main effects). In summary, general application of a stricter criterion does not appear to be responsible for the low hit rate in the familiarised group, nor was there any specific shift in criterion for only the familiarised odours.

Although the exact processes involved are unclear, the key finding of this study is that there is a disruptive effect of experimental familiarisation on n-back performance. Importantly, that there was this disruptive effect should be considered independently of any advantage to working memory performance from the use of high verbalisability odours. Indeed, such findings may support the proposal in Experiment 3 that conceptual saliency is responsible for a working memory performance advantage (e.g. by enabling the successful binding between item and context, Oberauer, 2005; Rajaram, 1998). This is because the advantage for verbalisable odours in working memory is independent to any effects (in this case, disruption to performance) of perceptual experience through familiarisation.

An alternative suggestion, however, may be that the familiarity gained from a familiarisation task is different to the normative familiarity observed for the high verbalisability set. Although Wilson and Stevenson (2006) would argue that normative familiarity and familiarity across short intervals are the consequence of the same residual activation in memory, other proposals suggest that whilst familiarity in working memory arises from the continued activation of items during the task, familiarity in long-term memory arises from the degree of intra-item associations (Göthe & Oberauer, 2008; Mandler, 1980). However, familiarisation from a task performed 24-hours prior to the working memory test seems more likely to reflect normative familiarity than within-task familiarity. Instead, one might suggest the disruptive effect of pre-experimental familiarisation occurred due to the nature of the familiarisation task. That is, experiencing odours in an experimental setting contrasts the normative familiarity and semantic associations that would typically be gained from a real-world setting (e.g. Degel et al., 2001).

In summary, the present study demonstrates an important role of item familiarity in working memory performance, though rather than facilitation through enhanced discriminability; this effect was disruptive to performance in the n-back task. It should be noted that the present study may not have included sufficient exposures for participants to gain the familiarisation advantage observed in typical familiarisation, or as noted above may not have replicated the typical experience of odour learning that occurs in non-laboratory settings (cf. Sinding et al., 2015 for a familiarisation effect on odour perceptual using a similar preliminary task). Future research should also assess the disruption of familiarisation under a blocked design, which one might suggest would still occur for familiarised odours, but is unlikely to impact performance for high familiarity odours that were not previously exposed. For completeness, a pre-exposure group to only high familiarity odours could also be included to determine whether this disruption can be invoked on odours with both high and low normative familiarities.

3.5 General Discussion

The goal of the current chapter was to apply the normative data produced in Chapter 2 to replicate the previous finding of a working memory advantage for odours that are verbalisable (Jönsson et al., 2011), and to explore this advantage in relation to strategies employed in the n-back task.

The observed advantage in the n-back task for verbalisable odours found across Experiments 1-4 serves to validate the normative data collected in Chapter 2. The normative data were used to separate odour sets on verbalisability in all of the above experiments. The replication of a verbalisability advantage (Jönsson et al., 2011) supports conclusions that verbal labelling, or some covariate to verbal labelling ability, facilitates working memory performance in the n-back task. Although the interaction in Experiment 1 lent support to the proposal that verbal labels are generated and refined throughout the task, this effect was not replicated in future experiments. Furthermore, whilst the advantage for verbalisable odours was robust throughout each replication of the task as measured by *A'* sensitivity, some variation across experiments were observed as to whether the effect was driven by a change in hit rate (Experiment 3 and 4), false alarms (Experiment 2), or both (Experiment 1).

Demonstration of olfactory working memory in the n-back task (with above chance sensitivity for low verbalisability odours) demonstrates that an odour representation can be re-activated and compared to the presented stimulus. However, n-back performance for both low and high verbalisability odours may be mediated by the ability to verbally re-code the items that subsequently allows rehearsal (Murphy et al., 1991). Furthermore, if verbal rehearsal is not occurring, it is debated whether such representations are consciously accessible and thus available to active maintenance and updating processes required in the n-back task (Arshamian & Larsson, 2014; Stevenson, 2009). Experiment

2 provides evidence that performance in the n-back task is not reliant on a verbal rehearsal process for odours either high or low on normative verbalisability. Specifically, reducing verbal rehearsal opportunities (through inter-trial CA) did not attenuate the memory advantage for the verbalisable odours, nor produce a general decline in performance. In addition, Experiment 2 showed some effect of concurrent rotation on false alarm rates, potentially suggesting that the retention interval was used to maintain stimuli through a refreshing process (Raye et al., 2007). These findings warrant further investigation to assess the executive resources utilised for maintaining and updating olfactory information in the task. The concurrent articulation findings however are consistent with suggestions that olfactory memory is not dependent on verbal working memory processes (Andrade & Donaldson, 2007). These findings expand upon previous results by demonstrating a lack of verbal rehearsal during a task that ostensibly has both maintenance and updating requirements.

However, the n-back task, like other working memory tasks, can be influenced by automatic processes that assess the familiarity of items to make a task judgement (Loaiza et al. 2015). Indeed, an n-back decision may be made based on a familiarity signal and thus have no requirement for rehearsal in working memory, either through acceptance or rejection based on a strength criterion or through a low-control process where familiarity is compared to a signal-strength estimate for a target item (Juvina & Taatgen, 2007). In Experiment 3, the reliance on familiarity-based judgements or controlled strategies (recollection-based judgements) for completing the n-back task were assessed across odour verbalisability. The remember-know paradigm revealed a reduction in item recollection for low verbalisability odours with no corresponding increase in familiarity-based responding. This suggests that the advantage for high verbalisability odours is related to more successful application of a control strategy, and not a qualitative shift away from a control strategy towards either a low-control assessment of familiarity, or a criterion-based assessment of familiarity for low verbalisability odours. Instead, a large proportion of correct responses were the result of guess responses for low verbalisable odours. It was suggested that this is the result of reduced conceptual saliency, which affects the ability of participants to maintain and update bindings between the odour and its context in the task.

Finally, in Experiment 4 it was assessed whether the observed working memory advantage for high verbalisable odours was mediated by perceptual familiarity to the odours; based on previous studies relating high familiarity to fewer cognitive resources at encoding (Reder et al., 2015) and increased discriminability (Wilson & Stevenson, 2003a). Multiple presentations of odours in a preliminary task were applied to artificially induce item familiarity to previously unfamiliar odours. However, the experiment found no evidence to support the proposal that mere exposure can improve olfactory working memory, though the findings supported a disruptive effect of familiarisation, perhaps as a result of an unreliable familiarity signal. Consequently, although Experiment 4 did not reveal the expected findings related to perceptual familiarity, it does support proposals that item familiarity is important in an n-back task decision, and that the verbalisability advantage for odours occurs from something other than perceptual familiarity (e.g. conceptual saliency).

In summary, this chapter has replicated a verbalisability advantage for odours in the n-back task, but has presented evidence against this advantage being due to verbal rehearsal (Experiment 2), or perceptual familiarity (Experiment 4). The performance advantage instead appears to be due to an increase in the ability to maintain and update bindings between the odour and its context in working memory (Oberauer, 2005), though it should be noted that there was no evidence to support a shift in strategy

(Experiment 3). Indeed there was a general increase in false alarms in a concurrent rotation task that supports similar application of working memory resources in both verbalisability conditions (Experiment 2).

Chapter 4: An individual differences analysis of cross-modal n-back ability

4 Chapter summary

In this chapter an individual-differences methodology is applied, comparing individual performance in the olfactory n-back to other-modality n-back tasks. Specifically, the experiment examines the relationship between verbal, visual, verbalisable odour, and low verbalisability odour n-back performance. This approach is used to assess shared underlying processes across different tasks. Models of n-back performance describe a number of working memory processes that may be engaged in these tasks, including maintenance and updating of the n-back rehearsal window, and resolving interference from no-longer relevant items and from lure items with a strong familiarity signal (Chatham et al., 2011; Harbison et al., 2011; Juvina & Taatgen, 2007; Kane et al., 2007; Szmalec et al., 2011).

Participants performed 2-back tasks with low verbalisability odours, verbalisable odours, and abstract shapes; and a 3-back task with letter stimuli. Furthermore, participants performed an olfactory discriminability task to examine perceptual differences across odour sets. An index of working memory capacity was calculated from hits and false alarms, producing an *A'* score for each task per participant. A correlation matrix revealed moderate relationships between verbal, visual, and only verbalisable odour n-back performance. That is, low verbalisability odour performance was unrelated to the verbal and visual n-back tasks.

The findings support a similar executive demand in verbalisable odour working memory to verbal and visual working memory. This is discussed in terms of available semantic information, the additional coding of odours with verbal information (Paivio, 1990), and the conceptual salience gained when an odour is identified (Rajaram, 1998). Together, the findings have important implications regarding the ability to apply working memory resources to odours when available semantic information is low. It is suggested that internal attention to an olfactory representation is not available without the inclusion of this additional semantic information.

An individual differences analysis of cross-modal n-back ability

4.0 Chapter Introduction

4.0.1 Shared variance in measures of working memory capacity

The use of control strategies in the n-back task, and how the procedure might relate to other measures of working memory capacity, is discussed in detail in Chapter 3. It is generally agreed that indices of working memory capacity denote some use of attentional control to maintain and manipulate stimulus information, resolve interference, and perform other executive tasks (Baddeley & Hitch, 1974; Engle & Kane, 2004; Oberauer, 2009). Consequently, individual differences research will typically show a correlation between multiple measures of working memory capacity, which supports the application of this executive/attentional system for fulfilling the task requirements (Engle & Kane, 2004; Schmiedek, Hildebrandt, et al., 2009; Wilhelm et al., 2013). This modality independent process is also supported in neuroimaging, where prefrontal activations are organised according to the task performed rather than the stimulus type used (e.g. Owen, 1997).

However, whilst a verbal n-back task is used as a common measure of working memory capacity for neuroimaging research, the actual utility of this procedure as a measurement of working memory is complex and equivocal (see Chapter 3 for a full discussion). In a meta-analysis by Redick and Lindsey (2013), the n-back task was weakly correlated to complex span tasks, suggesting the two cannot be used inter-changeably in research applications. In contrast, Schmiedek et al. (2014) produced latent factors from performance in three complex span tasks (reading span, counting span, and rotation span) and from n-back task performance (using a numerical and spatial n-back), and these two factors correlated substantially. Their reasoning for commonly-found poor correlations between complex span and n-back tasks were due to paradigm-specific variance (e.g. the use of recall or recognition procedures), content-specific variance (e.g. the requirement for rapid counting in the counting span task, compared to visuo-spatial processing the rotation span task), and measurement error (e.g. ceiling and floor effects). This finding supports a hierarchical model of working memory performance where both n-back and complex span tasks involve the same higher-order construct in working memory, though of relevance for the present study is the finding that it is unsuitable to assess cross-modal relationships using cross-task comparisons. This is because the multiple sources of variance that result in weak relationships between tasks will mask whether the different-modality tasks applied here reflect the application of similar working memory resources.

Further support from Schmiedek et al. for the role of a higher-order working memory process that drives n-back performance is seen from the task's strong relationship with fluid intelligence (measured by Raven's Advances Progressive Matrices), which is proposed to occur because attentional control is essential for both skills (Carpenter, Just, & Shell, 1990; Jaeggi et al., 2008; Schmiedek, Hildebrandt, et al., 2009; cf. Wilhelm et al. 2013 for a binding explanation of working memory capacity). Indeed, transfer effects have been observed from training in the n-back task to measures of fluid intelligence, also attributed to the requirement in both tasks for control of attention (Jaeggi et al., 2008).

A general mechanism in the n-back task is supported by modality-independent brain regions implicated during the procedure. Owen, McMillan, Laird, and Bullmore (2005) assessed different-modality n-back tasks in a meta-analysis of functional neuroimaging studies. Included in the analysis were multiple verbal and non-verbal n-back tasks (e.g. shapes, faces, numbers, words, and fractals) that required either identity or spatial judgements. Their findings saw robust activation in the dorsolateral prefrontal cortex, associated with strategic control of working memory processing (i.e. frontal lobe damage has been associated with the application of inefficient strategies in working memory, Owen, Morris, Sahakian, & Polkey, 1996); and the ventrolateral prefrontal cortex, implicated in the mapping of stimuli to responses upon presentation of targets or non-targets (Andersen and Buneo, 2003 cited in Owen et al., 2005). Their analysis noted these prefrontal activations in the olfactory n-back task performed in Dade et al. (2001), though it has been discussed previously and in Jönsson el al. (2011) that these findings may be explained by verbal processes in the stored odour representation rather than perceptual representations per se. However, whilst there appeared to be evidence for amodal activation during the n-back, there are also findings that support modality-specific regions of activation in the n-back task. Owen et al. (2005) acknowledged hemispheric lateralisation in frontoparietal regions related to whether the stimuli was verbal or non-verbal. Furthermore, in an n-back imaging study Knops et al. (2006) showed activation in the horizontal intraparietal sulcus for numerical stimuli, that they attribute to processing of *averbal semantics* (i.e. assessment of magnitude). That is, it is proposed that information in a working memory task will not only be represented as a phonological code, but will benefit from additional processing of semantic information.

The relationship between performance levels on different modality n-back tasks is another method by which generalised processing can be examined, and, in general, this method reveals strong different-modality task relationships. For example, numerical and spatial n-back procedures have shown a strong relationship (r = .66 across accuracy measures, Schmiedek et al. 2014), though other comparisons, between visuospatial and auditory 2-back tasks, have revealed weaker correlations (r = .35 across accuracy measures, Jaeggi et al. 2010). This is of interest to the present task because a relationship between olfactory n-back performance and n-back performance from other modalities can elucidate the processes engaged in these tasks when performed with olfactory stimuli. This is because verbal and visual n-back tasks have previously shown a relationship not only with each other, but also to the ability to apply controlled working memory resources in latent variable studies (Schmiedek et al., 2014; Wilhelm et al., 2013). Specifically, this means that a relationship between olfactory n-back performance and performance in verbal and visual n-back tasks can be interpreted as a general application of controlled working memory processing. It should be noted, however, that analysis of individual differences across different-modality n-back tasks is naturally limited by the reliability of the measure itself (Jaeggi, Buschkuehl, et al., 2010; Redick & Lindsey, 2013). Performance in the n-back task is most reliable when $n > 1$ (Friedman et al., 2008; Jaeggi, Buschkuehl, et al., 2010; Kane et al., 2007; Shelton, Elliott, Hill, Calamia, & Gouvier, 2009), though in Jaeggi et al. (2010) the split-half reliability of 2-back tasks varied between $r = .26$ and $r = .85$. This led the authors to conclude that the n-back task is not a suitable tool for measuring participant individual difference. To be clear, if within-participant variance is high throughout the task, then cross-modality comparisons have less validity due to uncertainty as to whether the 2-back score is a true representation of ability. However, Redick and Lindsay (2013) suggest the opposite, concluding in their meta-analysis that the n-back task does produce acceptable reliabilities ($r > .70$) in several studies (e.g. Kane et al., 2007; Oberauer, 2005; Schmiedek, Hildebrandt, et al., 2009; N Unsworth & Spillers, 2010). This is also supported in Schmiedek et al. (2014), with reliability estimates of $\alpha = .92$ and $\alpha = .95$ for number and spatial n-backs, respectively. Taken together, the findings suggest that reliability is not problematic for an individual-differences assessment of n-back performance.

4.0.2 Semantic information in olfactory working memory

The previous chapter showed performance in an olfactory 2-back task that was better when odours were verbalisable (see also Jönsson et al., 2011) and that this advantage was not due to the use of verbal rehearsal strategies to maintain these odours (Experiment 3.2). Whilst there is equivocal evidence regarding the ability of participants to consciously access a stored representation of olfactory information that would allow some rehearsal and updating strategy to be performed (e.g. Arshamian & Larsson, 2014; Stevenson, 2009), the dual-task findings in Chapter 3 lend some support for application of resources for maintaining odour representations, where false alarms increased with an inter-trial rotation task (although this was not reflected in hits or *A'* sensitivity). This impairment did not interact with odour verbalisability, however, suggesting that any change to an odour representation with increased knowledge (e.g. verbalisation or identification; Yeshurun et al., 2008; Zelano et al., 2009) does not change the contribution of working memory processes for effective completion of the task (e.g. refreshing, Raye et al., 2007). In summary, these findings suggest that the working memory processes applied for maintaining odours are similar for both low and high verbalisability stimuli.

In Experiment 3.3 it was shown that an advantage for verbalisable odours was due to increased recollection of these verbalisable odours. In working memory, the measurement of recollection reflects the application of controlled working memory resources (Baddeley, 2012; Barrett, Tugade, & Engle, 2004; Engle & Kane, 2004; Loaiza et al., 2015). For example, participants demonstrate a selective decrease in recollection estimates with faster presentation rates, and estimates of recollection are more sensitive to changes in fluid intelligence (Loaiza et al. 2015). The reduction in recollection therefore suggests a reduced ability to apply these resources to odour representations when verbalisability is low. Together, these findings suggest a control strategy applied in the olfactory n-back task that maintains items in working memory, and updates the link between this item and its serial position as the task progresses. However, the recollection of item and its linked serial positions fails more often for odours that are difficult to verbalise.

When semantic information is mapped to perceptual information its conceptual saliency is increased, and this is linked to greater recollection (Rajaram, 1998). This can explain the observed similarities between odour recollection in identified odour memory and verbal memory (M. J. Olsson et al., 2009). That is, Olsson et al. (2009) demonstrated similar episodic recognition performance and recollective experience between identified odours and verbal memory, whilst unidentified odour memory showed lower levels of both. In working memory, available semantic information may therefore mediate the ability to maintain odours online (Jönsson et al., 2011) and is perhaps responsible for equivocal findings related to consciously accessible odour imagery (Stevenson et al., 2007; Tomiczek & Stevenson, 2009).

Yeshurun et al. (2008) suggest that unidentified odours are recognised by their olfactory pattern and additional poor quality verbal information, whilst identified odours are recognised by their olfactory pattern and a centrally-mediated representation that includes strong semantic and verbal information (R. A. Frank et al., 2011; Stevenson, Boakes, & Wilson, 2000). Similarly, Zelano et al. (2009) describe activation in both the piriform cortex and prefrontal language areas when performing a short-term odour memory task, but these activations show a double dissociation when odours are identified or unidentified. That is, whilst the piriform cortex is favoured for unidentified odours, the prefrontal areas are favoured for identified odours. Together, these differences in representation may lead to a shift in the way odours are processed in working memory. However, the findings in Chapter 3 are equivocal over whether the availability of semantic information results in a qualitative shift in the way odours are processed (that is, working memory resources used may change when semantic information is available). It is possible that the working memory processes are simply less reliable when a verbal memory trace is not available due to lower conceptual salience (Rajaram, 1998).

4.1 Individual differences in multi-modal n-back performance

This study applies an individual-differences approach to further examine whether available odour semantic information (through normative odour verbalisability) is related to the utilisation of controlled working memory resources. The experiment assesses the relationships between n-back performance using low verbalisability odours, high verbalisability odours, abstract shapes, and letter stimuli. Furthermore, individual differences in discriminability are also considered, as discriminability is a limiting factor in any measure of working memory performance (Jönsson et al., 2011).

A relationship is expected between visual and verbal n-back performance due to previously observed similarities in working memory performance across these modalities (e.g. Schmiedek et al., 2014). However, the prediction regarding relationships with odour n-back tasks are less clear. If performance in these multiple n-back tasks arises from a common working memory resource, the tasks should covary. The dual-task findings in Experiment 3.2 suggest that similar processes may be engaged for low verbalisability and verbalisable odours (though the evidence for this, and subsequent evidence from the remember-know task are equivocal on this), and would therefore predict this relationship between all n-back task modalities despite lower working memory performance for the low verbalisability odours. However, if the availability of semantic information mediates application of working memory control processes, a relationship between verbalisable odour working memory and the verbal and visual tasks is predicted, but with no relationship to low verbalisability odour memory. Finally, it is predicted that performance on both low and high verbalisability discrimination tasks will be related to olfactory performance in both low and high verbalisability n-back tasks, due to the perceptual processing requirements proposed for all versions of the n-back task.

4.1.1 Method

4.1.1.1 Participants

Fifty-six participants (44 females, 12 males, mean age = 23.91, *SD* = 6.64) were recruited from Bournemouth University as part of a course credit requirement. The same exclusion criteria from earlier chapters were applied. Ethical approval was gained for all aspects of the study through the Bournemouth University ethical procedures.

4.1.1.2 Materials

Olfactory stimuli. Fourteen olfactory stimuli were selected for use in this experiment (Table 3). Half of these were classified as low verbalisability, and half as high verbalisability, based on normative ratings in Chapter 2. Odour sets differed significantly on these verbalisability scores, $t(12) = 12.96$, $p < .001$, $d = 6.93$, $BF_{10} > 1{,}000$, and on familiarity scores, $t(12) = 22.22$, $p < .001$, $d = 11.88$, $BF_{10} > 1{,}000$. Furthermore, the odour sets were balanced on ratings of intensity, and Bayes Factor analysis revealed support for no difference in intensity, $t(12) = -0.19$, $p = .851$, $BF_{01} = 2.21$. Whilst pleasantness ratings did differ between the two sets, $t(12) = 8.00$, $p < .001$, $d = 4.28$, $BF_{10} > 1{,}000$, there was evidence for no difference in the hedonic strength across the two sets. This was calculated as the deviance from a neutral midpoint on the pleasantness rating scale (Chapter 2), $t(12) = 0.67$, $p = .515$, $BF_{01} = 1.93$. Responses were collected using a Cedrus Response Box, and recorded using Superlab 5 (Cedrus, 2015).

Table 3

Normative ratings and grouping of olfactory stimuli used in the low and high verbalisability odour n-back tasks.

Odour	Task	Verb.	Fam.	Int.	Pleas.	Hed. Str.
Lime*	High	2.73	5.70	5.06	5.16	1.40
Pear	High	2.62	5.82	5.16	4.40	1.40
Blackcurrant	High	2.44	5.67	4.85	5.48	1.73
Marzipan	High	2.73	6.12	5.27	4.96	1.65
Spearmint	High	2.71	5.90	4.96	5.08	1.48
Aniseed Balls	High	2.61	5.88	5.40	3.98	1.50
Sports Rub	High	2.69	5.60	5.52	4.21	1.08
Cheddar Cheese*	Low	1.24	3.14	5.27	2.35	1.86
Ginger	Low	1.66	3.39	5.22	3.10	1.39
Sea Shore	Low	1.53	2.96	5.20	2.20	1.84
Rum Barrel	Low	1.26	3.10	5.18	2.68	1.56
Carbolic Soap	Low	1.05	3.61	5.10	3.12	1.33
Patchouli	Low	1.62	3.55	5.06	3.02	1.31
Mouse	Low	1.53	3.36	5.06	2.70	1.50

* Buffer items not included in analysis

Visual stimuli. Seven irregularly-shaped polygons (Chuah, Maybery, & Fox, 2004) designed to prevent verbal rehearsal strategies (Attneave, Arnoult, & Attneave, 1956; Smith et al., 1995) were used as 2-back stimuli. These were presented using

Opensesame (Mathôt et al., 2012) in the centre of a 22-inch 60hz monitor as black line drawings on a white background, within a black border square of 62px by 62px (See Figure 10).

Verbal stimuli. Eight phonologically dissimilar consonants were selected (B, F, H, K, M, Q, R, X) (Kane et al., 2007), and displayed centrally on a 22-inch 60hz monitor in size 21pt. monospaced font. Stimuli were randomly presented in lower or upper case to limit responses based on the visual features of the letters (although one might argue that *f, k,* and *m* are visually similar in lower and upper case forms), and stimulus presentation timings and trial responses (Figure 10) were controlled by OpenSesame.

4.1.1.3 Design

Olfactory 2-back task. A continuous yes/no recognition paradigm was employed on two 26-item sequences, where each trial necessitated a judgement as to whether the item had been presented two items previous. The low and high verbalisability odours were presented in the same blocked design used in Experiments 3.1 and 3.2, and Jönsson et al. (2011, Experiment 2), meaning participants experienced 26 trials of one odour set followed by 26 trials of the other set. A unique low or high verbalisability odour, corresponding to the odour set being tested, was presented as a buffer item in the first two trials of each 2-back task, where a 'no' response was guaranteed. Each remaining odour appeared as a lure three times and as a target once. In addition, the presence of *close-lures* was increased compared to earlier experiments to more closely match the number of targets, with a task containing 5, or 6, n+1 and n-1 lures. This adjustment was important for two reasons. First, it discourages a reliance on a familiarity-based strategy simply because the payoff from using a recollection-based strategy was not worth the use of additional resources (Ralph, 2014). Furthermore, it allows analysis of responses based on only recent-lure and target decisions, meaning judgements that may be based solely on a familiarity criterion (non-recent lures) are not included in the index of working memory ability.

A blocked design allowed the two sets to be considered as independent memory tasks, minimising cross-contamination of items from the other odour set. That is, a criticism of Experiments 3.3 and 3.4 was that the odour presented in the n-1 position, if different across the dimension of interest to the 2-back item and probe, may have influenced the probe's acceptance or rejection through some combination of familiarity-based responding and a process-of-elimination strategy.

The orientation of targets and lures within tasks were identical for all participants, but were counterbalanced such that the half of the participants performed the low verbalisability task first, and the other half the high verbalisability task first. In addition, order of trials (targets and lures) was counterbalanced between the low and high verbalisability tasks. The former counterbalancing accounted for practice effects, whereas the latter for differences in trial order difficulty. That is, the slightly different number of close lures in a sequence may have affected the difficulty of the task, but this was balanced across tasks between participants.

Visual 2-back task. The visual 2-back task consisted of 2 blocks of 26 items, in an identical trial sequence to the two olfactory sequences. Two identical buffer images preceded the 24 critical 2-back trials in each block, and did not occur again in either sequence. The presentation order of the two visual blocks was randomised across participants.

Verbal 3-back task. Pilot work (data unavailable) suggested that a 2-back task with verbal stimuli was close to ceiling, meaning a 3-back task was instead selected for verbal stimuli. The task necessitated a judgement whether the currently presented item matched the item presented three trials previous. Consequently, the first three items in a block sequence were guaranteed 'no' responses, after which the fourth trial was compared to the first, fifth to the second, etc. Participants completed 3 unique blocks of 40 trials, and were given the opportunity for a break between each block. For each sequence, a letter appeared once as a target, and four times as a lure, totalling 8 targets and 32 lures in each block. Furthermore, a minimum of 7 *n*+1, *n*-1, and *n*-2 lures occurred within a sequence, and a maximum of 10. The trial order was pre-generated and the same for all participants. It should be noted that this task therefore differs substantially to the 2-back tasks, particularly in terms of general working memory load and the presence of recent-lure trials.

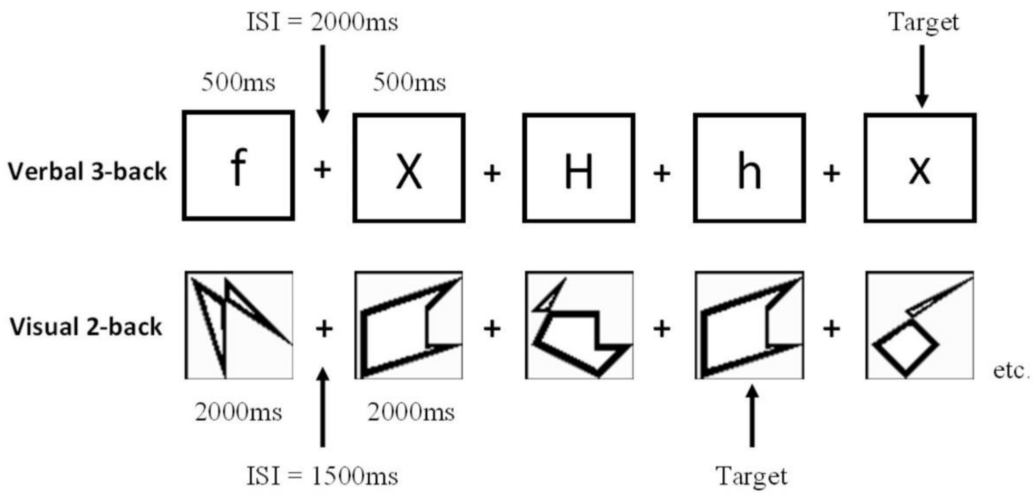

Figure 10. Schematic diagram of verbal 3-back and visual 2-back tasks.

Olfactory discriminability task. A paired discrimination task was employed where participants made a *same/different* judgement for two odours presented in succession. There were 42 possible non-match combinations, where an odour within a set was paired with every other odour in that set (i.e. a high verbalisability odour was only tested against other high verbalisability odours). There were 14 targets, with each odour appearing a matched pair once. Due to the large number of discriminability comparisons (56), participants performed half of the possible comparisons. To be clear, each participant performed 28 trials containing 14 verbalisable and 14 low verbalisability discrimination trials, consisting of 7 targets and 21 lures. Participants were randomly allocated to the first or second combination of odour pairs, and the presentation order of these pairs was randomised for each participant.

4.1.1.4 Procedure

Written consent was gained from all participants. Testing took place in a well-ventilated laboratory at Bournemouth University, and participants completed three versions of the n-back task and an odorant discriminability test in a single session, lasting approximately 50 minutes. Participants performed the computer-based (verbal and visual) and olfactory n-back tasks in a partially-counterbalanced order. Specifically, participants could perform either the olfactory or computer-based tasks first, but would always complete the two computer-based tasks together (though these two tasks were themselves performed in a counterbalanced order). The odour discriminability task always followed the n-back procedures, performed after a 10 minute break from the end of whichever n-back task was performed last.

Olfactory 2-back task. Participants sat opposite the experimenter, separated by a wooden screen. An 8-item visual stimuli practice task, identical to the practice trials in Chapter 3, was performed to familiarise participants with the task even if they had already completed the computer-based 2-back task. Odours were presented birhinally by the experimenter holding the odour under the nose of the participant for 2 seconds. An 8-second ISI separated presentation of the odours, during which participants made a yes/no 2-back decision. Participants were given water to sip throughout the task, and a break between the low and high odour sets to prevent olfactory fatigue.

Computer-based tasks. Computer-based verbal and visual n-back tasks were performed with participants sat approximately 50cm in front of a computer monitor. The software presented instructions to participants about the task ahead, and directed whether the task demand was to perform 2-back or 3-back comparisons. Participants performed a 10-item practice task for the modality they were about to be tested on, which the researcher monitored to ensure participants understood the requirements of the task. Participants were allowed to retake the practice trials until the researcher was satisfied with their understanding, though no participant required more than two completions of the 10-item sequence.

The verbal 3-back task presented participants with a sequence of letters separated by a fixation cross. Letter stimuli were presented for 500ms, followed by a fixation cross for 2000ms, based on similar timings from the 3-back task applied in Jaeggi et al. (2010). Judgements could be made during either the stimulus presentation or fixation, and were made using the 1 key (No match), and the 3 key (Match). Presentation of the next item was not dependent upon receiving a participant response, and a missed response was logged as incorrect.

The visual 2-back task presented participants with a sequence of abstract polygons, each displayed for 2000ms, separated by a 2000ms fixation cross. The presentation time of the visual stimuli was slowed compared to the verbal stimuli because pilot data indicated a 500ms presentation was too fast for effective completion of the visual task. The response procedure followed that described for the verbal task, except participants were required to match the present trial with that seen two items previous.

4.1.2 Results

A different approach to analysis was taken in the current study compared to the n-back tasks in Chapter 3. Specifically, greater control of close-lures in the current n-back task sequences allowed assessment of false alarms for only these challenging trial types. These lures are suggested to be more involved in controlled responding in working memory due to the need to resolve conflicting evidence between a within-experiment familiarity signal and recollection process (Harbison et al., 2011; cf. Kane et al., 2007 for the use of close lures as an index of working memory ability that is poorly correlated to complex span). The measurement of only close-lure false alarms produces an *A'* score that is not inflated by easy non-recent lure rejections, and provides a more sensitive assessment of differences in working memory ability at the expense of using a lower number of trials (Ralph, 2014).

4.1.2.1 N-back ability

A' sensitivity. Above-chance performance was assessed using one sample t-tests against an *A'* score of 0.5. All comparisons were significant (*ps.* $< .001$), and the data strongly supported above-chance performance, $BF_{10} > 1{,}000$.

Figure 11 shows the mean *A'* sensitivity score across the four n-back tasks. *A'* scores were entered into a within-participants ANOVA across the four tasks, though it should be noted that such comparisons are problematic given the differences in methodology across modalities. The findings can, however, be used as a general indicator of how well participants were able to perform each task. A significant main effect of task modality was found, $F(3, 165) = 10.58$, $p < .001$, $\eta_p^2 = .16$ (and a main effects model preferred to the null model, $BF_{10} > 1{,}000$). This main effect was further analysed with paired comparisons[2] and Bayes Factor t-tests, and revealed support for better performance in the verbalisable odour task than the low verbalisability odour task (p = .005, BF_{10} = 13.27). This replicates the verbalisable advantage reported in Experiments 3.1-3.4 using a more stringent calculation of *A'* sensitivity. There was also support for a difference between visual 2-back and low verbalisability odour 2-back performance (p = .003, BF_{10} = 11.59), and between visual 2-back and verbal 3-back performance (p < .001, BF_{10} > 1,000), with better performance in the visual 2-back task. Verbalisable odour performance was also higher than the verbal 3-back performance, with strong evidence for a difference between the two scores (p < .001, BF_{10} = 753.98). Finally, there was evidence against a difference between verbalisable odour performance and visual performance (p = .657, BF_{10} = 0.16), and between low verbalisability odour and verbal 3-back performance (p = .334, BF_{10} = 0.23).

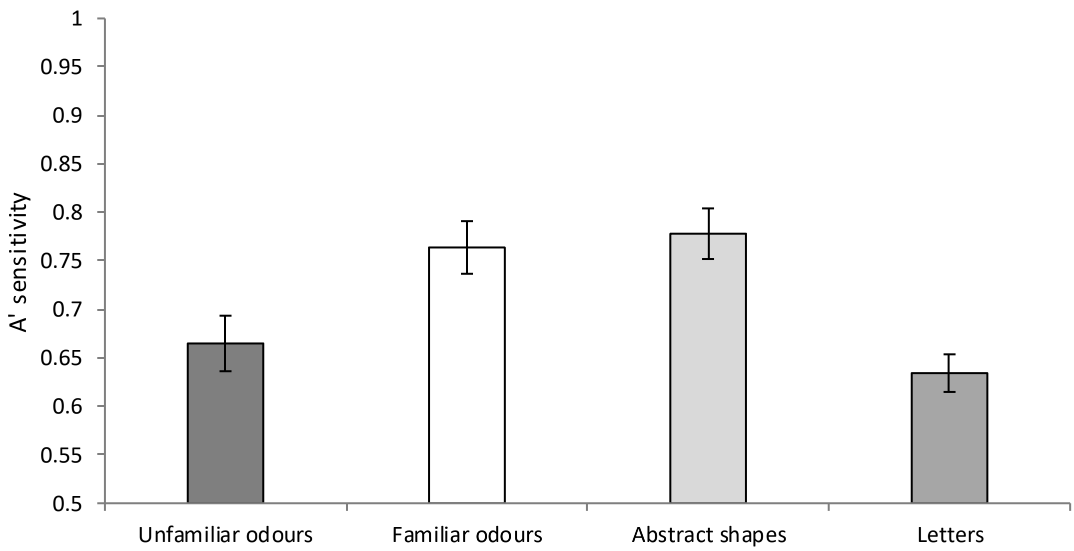

Figure 11. Mean *A'* sensitivity scores across the four different modality n-back tasks.

[2] The p values reported in this section are uncorrected, but should be compared to α = .017.

Hit and close-lure false-alarm rate analysis. The differences in hit and false alarm rates were broadly similar to the *A'* comparisons. That is, there was similar performance between verbalisable odour and visual n-back tasks, and better performance than the verbal 3-back task. This did, however, include higher hit rates in the low verbalisable odour n-back than in the verbal 3-back task, despite evidence for no difference between the two tasks in *A'* scores (see Appendix D for a full write-up of these results).

4.1.2.2 Odorant discriminability

The discriminability of low verbalisability and verbalisable odours was assessed to explore the extent of the advantage odour verbalisability has on item discriminability. *A'* was calculated from the proportion of hits (correct *match* responses) and false alarms (incorrect *match* responses), and compared using a paired t-test and Bayes Factor analysis. This revealed evidence against a difference between low and high verbalisable odour discriminability, $t(55) = 0.96$, $p = .341$, $d = 0.19$, $BF_{10} = 0.23$, despite the performance differences observed in the n-back task, above. Furthermore, *A'* sensitivity was significantly above chance for low, $t(55) = 66.11$, $p < .001$, $d = 8.84$, $BF_{10} > 1{,}000$, and high verbalisability odours, $t(55) = 42.15$, $p < .001$, $d = 5.63$, $BF_{10} > 1{,}000$.

4.1.2.3 Correlation matrix

Table 4 shows a correlation matrix computed for *A'* scores across tasks. First, the theoretically interesting comparisons between odorant discriminability and olfactory working memory ability showed a positive relationship between low verbalisability odour discriminability and both low verbalisability ($r = .30$, $BF_{10} = 4.11$) and verbalisable ($r = .32$, $BF_{10} = 5.09$) odour working memory. In contrast, there was evidence against a positive relationship between verbalisable odour discriminability and low ($r = -.05$, $BF_{10} = 0.13$) and high ($r = -08$, $BF_{10} = 0.29$) verbalisability odour working memory. This is a perplexing pattern of results which suggest a qualitative difference between verbalisable odour discriminability and low verbalisable odour discriminability, but also from both n-back tasks regardless of odour verbalisability.

Comparisons between n-back performance revealed anecdotal (*BF* < 3, Jeffreys, 1961) support for a moderate positive correlation between low and high verbalisability odour working memory, $r = .27$, $p = .047$, $BF_{10} = 2.24$. In contrast, there was anecdotal evidence against a positive correlation between low verbalisability odours and visual working memory, $r = .14$, $p = .292$, $BF_{10} = 0.49$, and verbal working memory, $r = .19$, $p = .168$, $BF_{10} = 0.77$. For verbalisable odour working memory, however, there was support for a moderate correlation with visual working memory, $r = .30$, $p = .027$, $BF_{10} = 3.52$, and with verbal working memory, $r = .30$, $p = .026$, $BF_{10} = 3.70$. Finally, there was strong support for a moderate positive correlation between verbal and visual working memory, $r = .49$, $p < .001$, $BF_{10} = 371.28$.

Table 4

Correlation matrix of A' scores for the four n-back tasks, and two tests of odour discriminability.

	1.	2.	3.	4.	5.	6.
1. Low verbalisability odour 2-back	—	.27*	.14	.19	.30*	-.05
2. Verbalisable odour 2-back		—	.30*	.30*	.32*	.08
3. Visual 2-back			—	.49**	.07	.03
4. Verbal 3-back				—	-.06	.09
5. Low verb. odour discriminability					—	-.05
6. Verbalisable odour discriminability						—

* $p < .05$

** $p < .001$

4.1.3 Discussion

The aim of the present study was to inspect the relationships between performances across different-modality n-back tasks. Although this design does not allow assessment of the precise strategies involved, it does allow differences in the underlying processes engaged across different stimuli to be examined. That is, an individual differences design would predict a participant to show similar n-back performance (relative to other participant scores) across tasks if similar working memory processes are being engaged. In addition, this experiment addressed an important methodological issue arising from earlier n-back studies in the thesis. Namely, recent lures were controlled and only included in the calculation of *A'*. This control ensures that participants cannot perform the task above chance using a strategy based uniquely on familiarity (as recent lures are more familiar than targets). Despite this control, the present experiment replicates the verbalisable advantage reported across the Chapter 3 experiments. Furthermore, it is important to note that n-back performance for low verbalisable odours remained above chance. This shows that n-back performance for low verbalisable odours was not supported uniquely by a reliance on familiarity.

The results of the correlational analysis showed an expected relationship between the visual 2-back and verbal 3-back task which supports a general executive requirement when performing the n-back procedure (e.g. Schmiedek, Hildebrandt, et al., 2009; Schmiedek et al., 2014; Wilhelm et al., 2013). Although weaker than the visual-verbal relationship, there were also moderate correlations between the verbalisable odour 2-back task and both verbal and visual tasks. This supports a shared mechanism for performing these n-back tasks, likely related to the application of controlled attention in working memory (Engle & Kane, 2004; Oberauer et al., 2000; Schmiedek, Hildebrandt, et al., 2009; Schmiedek et al., 2014; Wilhelm et al., 2013). Importantly, the low verbalisability odour 2-back task did not correlate with the verbal and visual n-back tasks, though there was (anecdotal) evidence for a relationship with the verbalisable odour n-back task.

Taken together, the evidence supports a working memory ability for verbalisable odours which engages shared resources to those in verbal and visual working memory. Such a finding should be considered in the context of an odour representation that contains multiple memory traces (e.g. Baddeley, 2000; Paivio, 1990), and the equivocal findings related to olfactory imagery and rehearsal (Arshamian et al., 2008; A. J. Johnson & Miles, 2009; Stevenson, 2009; Stevenson et al., 2007). That is, the verbalisable odours may be characterised by their available semantic information, and it this information that enables working memory resources to access a stored representation in memory (Tomiczek & Stevenson, 2009). This is consistent with the episodic memory findings in Olsson et al. (2009) where identified odours showed similar recognition ability and recollective experience to verbal stimuli. Like their findings, the data support that the mapping of semantic information to sensory stimulation produces odour memory with similar characteristics to verbal memory. Furthermore, like Olsson et al., the relationship between verbalisable odour working memory and the visual and verbal n-back tasks can be explained by increased conceptual salience (Rajaram, 1998), which is expected to facilitate the ability to maintain and update bindings in working memory (Oberauer, 2005).

The weak relationship between low verbalisability and verbalisable odour working memory suggests shared variance unrelated to the shared variance between verbalisable odour, verbal, and visual working memory. It is expected that such similarity, if not related to the application of working memory resources (and the lack of relationship between low verbalisability odours and the visual and verbal n-back tasks suggest that it is not), is instead related to individual differences in olfactory discrimination ability. However, why verbalisable odour discriminability was unrelated to the odour working memory tasks is unclear. Indeed, there was no evidence for a positive relationship between the two discriminability tasks, which is surprising because even if identified odours are processed differently to unidentified odours, there is still the need for olfactory perceptual processing (i.e. a pattern-matching process) before identification can occur (Stevenson & Wilson, 2007). Consequently, it suggests there may be different strategies applied in low and high verbalisable odour discrimination tasks, but that whatever strategy is applied in the verbalisable odour discrimination task is not related to the strategy adopted in the verbalisable odour n-back task. In addition, high performance in the discriminability task, which might suggest ceiling effects, do not appear to be problematic. This is because discriminability in the low verbalisability task (which did show a relationship with n-back tasks) was also high, and there was evidence against a discriminability difference between the two. Further research is therefore required to explore why this qualitative difference occurs for performance in only the verbalisable odour discrimination task.

It should also be noted that these discriminability findings contradict those in Jönsson et al. (2011), who showed better discriminability performance for their verbalisable odours. This cross-study disparity could be due to differences in the respective stimulus sets; however, despite this disparity in discriminability findings, the n-back findings replicate the previous experiments in Jönsson et al. (2011), and the earlier demonstrations of olfactory n-back performance in this thesis. Importantly, this verbalisability advantage has been demonstrated with a more sensitive assessment of the different trial types available in the n-back task, supporting above-chance performance that involves some control process to differentiate recently-present lures from targets (though as discussed in Chapter 3, strategies may still make use of a familiarity signal to make this decision, Juvina & Taatgen, 2007).

The aim of an individual-differences assessment of only n-back performance was to remove paradigm-specific variance that would be present in comparisons with other tests of working memory capacity, such as complex span (Schmiedek et al., 2014). Although a relationship was found between verbal n-back, verbalisable odour n-back, and visual n-back performance memory, some non-trivial differences between tasks should be noted. First, whilst $n > 1$ is generally considered a task requiring maintenance + manipulation (Ragland et al., 2002), the requirements for verbal stimuli was a 3-back comparison which not only adds difficulty, but may change the way targets and close-lures interact. For example, a disruptive effect to targets from immediately preceding *n*-1 lures occurs in a 3-back task, but not in a 2-back task (Kane et al., 2007). However, notwithstanding this methodological difference, relationships were found between the verbal 3-back and both the 2-back tasks for verbalisable odours and visual stimuli. Second, speeded presentation of visual and verbal information was implemented to prevent performance levels at ceiling, but this could have limited the rehearsal processes that could be engaged in a slower n-back procedure. The olfactory n-back task in comparison, with its 8-second ISI, may have encouraged a slower control process for completion of the task (Ralph, 2014).

Furthermore, the changing order of n-back tasks resulted in some participants performing the discriminability task immediately following the odour n-back task, in contrast to participants who performed the olfactory n-back task first, and received a larger break between olfactory-related tasks as a result. Although the closeness of olfactory tasks may have resulted in reduced discrimination ability for some participants, counterbalancing of task order should mitigate any particular bias emerging. In addition, there was a 10-minute break enforced after the end of the final n-back task, designed to mitigate fatigue effects from these heavy olfactory requirements.

In summary, the present study reports a relationship between n-back performance for verbalisable odour, verbal, and visual stimuli and is interpreted as support for application of controlled attentional processes for these stimulus types. In contrast, low verbalisability odours did not correlate with visual and verbal stimuli suggesting that such controlled attentional processes are not employed for low verbalisability odour working memory. The findings have important implications for the role of semantic information in olfactory memory, and the ability to engage internal attention to olfactory perceptual representations in memory. Specifically, the findings suggest that the ability to consciously access an internal representation of an olfactory experience may not be possible unless the representation is accompanied by semantic information.

Chapter 5. Proactive Interference in Olfactory Working Memory

5 Chapter summary

Chapters 3 and 4 have demonstrated working memory ability in olfaction that may be qualitatively different when items are verbalised, or when semantic information is available. One candidate explanation for this difference concerns differences in conflict resolution between high and low verbalisable odours for previously presented odours. To examine this proposition, Chapter 5 examines item-specific olfactory proactive interference (PI) effects for high and low verbalisable odours, and undertakes comparisons with verbal and non-verbal visual stimuli. Proactive interference in olfaction has been proposed to be particularly strong (Lawless & Engen, 1977), and has been proposed as evidence for independent processing of olfactory information in working memory.

Using a sequential recent probes task, no evidence for PI was found with hard-to-verbalise odours (Experiment 1). However, verbalisable odours did exhibit PI effects (Experiment 2). These findings occurred despite above-chance performance and similar serial position functions across both tasks. Experiments 3 and 4 applied words and faces, respectively, to the modified procedure, and showed that methodological differences cannot explain the null finding in Experiment 1. The extent to which odours exhibit analogous PI effects to that of other modalities is therefore argued to be contingent on the characteristics of the odours employed.

5.0 Chapter Introduction

Memory for a stimulus can be affected by stimuli that both precede (proactive interference: PI) and succeed (retroactive interference: RI) the to-be-remembered (TBR) stimulus. Whilst effects of interference on verbal stimuli are well established (Craig, Berman, Jonides, & Lustig, 2013; Jonides & Nee, 2006; Monsell, 1978; Postman & Underwood, 1973), the role of interference in olfactory memory is both under-researched and contradictory. Early work reporting a flat forgetting function for olfactory stimuli over extended retention intervals (Engen et al., 1973; Engen & Ross, 1973; F. N. Jones et al., 1978; Lawless & Cain, 1975) was attributed to weak RI coupled with strong PI (Lawless & Engen, 1977). However, this differential weighting in the levels of RI and PI contradict the serial position functions typically reported for olfactory stimuli with odours. That is, strong PI should produce a serial position function with primacy and weak recency. Specifically, monotonically increasing PI throughout the sequence should impair memory for latter list items to a greater extent than early list items. In direct contradiction to that prediction, primacy is rarely observed for olfactory short-term memory tasks (A. J. Johnson et al., 2013; A. J. Johnson & Miles, 2007, 2009; Miles & Hodder, 2005 c.f. Miles & Jenkins, 2000; Reed, 2000). Indeed, the presence of recency but not primacy (Johnson, Cauchi, & Miles, 2013; Johnson & Miles, 2007; Miles & Hodder, 2005; c.f. Miles & Jenkins, 2000; Reed, 2000), can be interpreted as evidence for RI, with RI monotonically decreasing throughout the sequence (it should be noted that support for RI in olfactory memory can also be found in Walk & Johns, 1984, and Köster, Degel, & Piper, 2002). Serial position data suggests that olfactory STM is not susceptible to PI and, the present set of experiments, therefore, seek to directly examine this proposition.

There is a paucity of studies examining PI in olfactory memory; with some support for PI found indirectly. For example, Valentin, Dacremont, and Cayeux (2011) showed recognition memory for odours declined as a function of experimental stage. Whilst this decline in performance may be interpreted as a build-up of PI, it is difficult to deconfound from the more general effects of olfactory fatigue (as reported by Reed, 2000). Köster et al. (2002) examined PI effects more directly using an implicit memory procedure in which two different experimental rooms were paired with odours. At the end of the study, participants were required to rate the extent to which certain odours 'fit' 12 different environmental contexts shown on a screen (of which 2 were the rooms used previously). They showed that the paired association for the second room-odour association can be disrupted (as indexed by a reduction in mean rating of 'fit') by memory for the initial room-odour association; a demonstration of proactive interference.

The distinction between Valentin et al. (2011) and Köster et al. (2002) is important because it highlights that one can subdivide PI effects into item-nonspecific and item-specific PI (Postle & Brush, 2004; Postle, Brush, & Nick, 2004). Non-specific PI can be conceptualised as a general build-up of interference following repeated exposure to a particular stimulus type and, thus, is difficult to differentiate from olfactory fatigue. In contrast, item-specific interference (e.g. Jonides & Nee, 2006; Monsell, 1978) concerns memory for a previously presented item (e.g. initial presentation of "lavender") interfering with a subsequent memory for that item (e.g. later presentation of "lavender"). Item-specific PI may, therefore, be taken as a direct measure of PI and is the focus of the present set of experiments.

The present study employs the recent-probes task (an established measure of item-specific PI, see Jonides & Nee, 2006, for review). In this task, participants undertake a series of trials, where each trial comprises a TBR memory set that typically numbers 4-items, and is followed by a single yes/no recognition probe. This probe is taken either from the preceding memory set (positive probe), or from an earlier TBR memory set (negative probe). The important manipulation for this task concerns previous exposure to the negative probe (also referred to as the 'lure'). Negative probes are divided into recent negative (RN) and non-recent negative (NRN). For the RN probes, the probe is taken from the memory set immediately preceding the current trial, whilst for the NRN probes, the probe is taken from the memory set presented 3 trials earlier. Thus, the key manipulation is the recency of the previous presentation of the negative probe. Item-specific PI effects are evidenced by both lengthened reaction times and an increase in errors for the RN probes in comparison to the NRN probes (Monsell, 1978). Both effects are typically interpreted via an increased need to resolve interference (Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998). That is, due to the strong memory for the RN probe, the individual experiences difficulty in determining whether that strong memory is a consequence of that item being included in the present trial, or being presented in a recent trial. This confusion regarding the origins of the RN probe is also reflected in metacognitive measures. Specifically, confidence ratings for correct rejections of the RN probe are typically lower than those for the NRN probe, and false alarms may reflect the presence of high-confidence intrusion errors (Jacoby, Wahlheim, Rhodes, Daniels, & Rogers, 2010; Wahlheim & Jacoby, 2011).

Different explanations have been proposed to account for these RN probe PI effects. Familiarity-inhibition (Jonides, Badre, Curtis, Thompson-Schill, & Smith, 2002; Mecklinger, Weber, Gunter, & Engle, 2003) states that the RN probe provokes powerful familiarity and this is typically associated with a positive response. Therefore, a correct rejection of the probe necessitates inhibition of that familiarity signal. It is this conflict between familiarity for the RN probe and memory for the TBR items in the present trial that causes errors and lengthened response times (Badre & Wagner, 2005). Similarly, context-retrieval models propose that the RN probe has less accuracy due to errors in the source memory for that item (Badre & Wagner, 2005). Put simply, the recency of the RN probe increases the likelihood that participants confuse the origins of the probe and believe that it was experienced in the present trial as opposed to the preceding trial.

The present set of experiments examines recent probe PI effects for olfactory stimuli. To date, the extent to which the recent probe PI effect is found across different stimulus types is equivocal (Jonides & Nee, 2006). For example, behavioural recent probe PI effects have been found with abstract symbols, letters, spatial locations, and unfamiliar faces (Badre & Wagner, 2005; Leung & Zhang, 2004; Mecklinger et al., 2003; Postle et al., 2004; Prabhakaran & Thompson-Schill, 2011), but not for colours and some shapes (Postle et al., 2004). It is not clear why these differences occur, but one explanation concerns stimulus distinctiveness (Mecklinger et al., 2003). According to Mecklinger et al. (2003), when stimuli are more distinct the memory representation for the items are more defined (and less fuzzy), resulting in a stronger match/mapping between the RN probe and previous presentation of that item. This stronger match results in an increased PI effect for the RN probe.

A stimulus characteristic pertinent for olfactory stimuli, that may mediate distinctiveness and consequently the recent probe PI effect, is verbalisability. Mecklinger et al. (2003) suggest stimuli that can be represented verbally are more distinct than non-verbal stimuli. This leads to the prediction that stronger PI effects should be observed for those olfactory stimuli that can be easily verbalised. However, previous studies have shown that manipulating levels of verbal facilitation does not affect the recent probe PI effect (Brandon, Jha, Trueswell, Barde, & Thompson-schill, 2003; Brandon, 2004, in Jonides & Nee, 2006). Indeed, inhibiting verbal labelling via a concurrent articulation condition has been shown to increase the effects of probe recency (Atkins, Berman, Reuter-Lorenz, Lewis, & Jonides, 2011). The Mecklinger et al. and Atkins et al. studies therefore provide competing predictions for the examination of olfactory PI effects and these competing predictions are directly tested in Experiments 1 and 2. According to the Mecklinger et al. (2003) account, one might predict stronger PI effects for verbalisable odours, compared to hard-to-verbalise odours, due to higher levels of distinctiveness when using verbal representations. In contrast, increased interference during articulatory suppression suggests that verbalisation may be a protective factor against recent probe PI effects (Atkins et al., 2011), and therefore one might predict less PI for verbalisable odours.

Experiment 1 directly tests short-term item-specific PI effects for olfactory memory using the recent probes task. Hard-to-verbalise odours are initially employed in order to investigate memory for olfactory percepts (rather than verbal labels of those percepts). Indeed, there is evidence that verbalisable and non- verbalisable odours may be represented differently in memory (e.g. Zelano, Montag, Khan, & Sobel, 2009), with the ability to verbalise odours shown to produce memory effects similar to that shown with words (Olsson, Lundgren, Soares, & Johansson, 2009). To be clear, since verbalisable odours may exhibit PI effects resulting from verbal rather than olfactory perceptions, Experiment 1 employs hard-to-verbalise odours.

In the classical recent probes procedure described above (see Jonides & Nee, 2006), participants are simultaneously presented with an array of TBR items (typically four) and followed by a single test probe. However, since odours cannot easily be differentiated following simultaneous presentation, a modification of the recent probes task is described in which the TBR items are presented sequentially. In this procedure, participants are presented with sequences of odours followed at test by a single yes/no recognition probe. The present study additionally includes an analysis of serial position for two reasons. First, this analysis provides another approach through which cross-modal STM comparisons can be made. There is some debate as to whether odours produce qualitatively different serial position functions to that of other stimulus types (e.g. see Johnson & Miles, 2009; Reed, 2000) and this study assesses cross-modal differences in yes/no recognition functions. Second, immediate yes/no recognition has been shown to produce a specific serial position function for visual stimuli (Hay et al., 2007; A. J. Johnson, Volp, & Miles, 2014; Kerr, Avons, & Ward, 1999); which in later studies using visual stimuli (Experiments 3 and 4) can be used to check whether the task has produced this canonical function. Based upon these previous studies (Hay et al., 2007; Johnson et al., 2014; Kerr et al., 1999), it is predicted that recognition for positive probes (i.e. odours that were presented in the preceding sequence) will produce serial position functions comprising recency but not primacy.

It is, however, responses to the negative probes that are important in determining any recent probe PI effects. Error rates (false alarms) are therefore compared for the RN probe compared to the NRN probe, with PI evidenced by increased false alarms for the RN probe. In addition, response confidence resolution is reported for the test probes as a more subtle measure of PI. Confidence resolution is an item-level gamma correlational measure (Roediger III & DeSoto, 2014) showing the intuitive positive relationship between confidence and accuracy (Wahlheim & Jacoby, 2011). However, one might predict a reduction in the strength of this correlation when responding to a RN probe (Brewer & Sampaio, 2012; Roediger III & DeSoto, 2014; Wahlheim & Jacoby, 2011). Since the strength of familiarity for the RN test probe is of less utility in accurately determining if that item was included in the present TBR memory set, the strength of the correlation between confidence and accuracy should decrease. Consequently, a recent probe PI effect for hard-to-verbalise olfactory stimuli would be evidenced by two outcomes. First, there should be an increased level of false alarms for the RN probe compared to the NRN probe. Second, there should be a reduced accuracy-confidence correlation for the RN probe compared to the NRN probe.

As stated previously, if distinctiveness determines recent probe PI effects (Mecklinger et al., 2003) a lack of PI is predicted for the hard-to-verbalise odours. However, if limiting verbal coding accentuates the recent probe PI effect (see Atkins et al., 2011), a strong PI effect for the hard-to-verbalise odours is predicted. Indeed, since parallels have been suggested between the processing of faces and odours (Kärnekull et al., 2015) a drop in accuracy is predicted for the RN probe that is comparable to that found with faces (Brandon et al., 2003 reported a 15% drop in accuracy for face stimuli).

5.1 Experiment 1: Recent-probes task with low verbalisability odours

5.1.1 Method

5.1.1.1 Participants

Twenty-four Bournemouth University students (non-smokers, mean age = 25.08, *SD* = 5.90, 11 females, 13 males) participated and received course credit or a £10 honorarium. Participants who self-reported olfactory impairments (e.g. symptoms of cold) were excluded. Additionally, participants aged older than 40 years were excluded (due to indications that olfactory-related abilities peak between the third and fifth decade, Doty et al., 1984). The study received ethical approval via the Bournemouth University research ethics procedure.

5.1.1.2 Materials

Olfactory stimuli. The experimental stimuli comprised one-hundred and sixty food and non-food related odorants selected from a corpus of 200 odorants, prepared by Dale Air Ltd. (www.daleair.com). Each odorant comprised 5ml of an oil-based liquid, and was stored in an opaque test tube in order to mask the odorant's colour. Twenty odours (Appendix E) were selected as the negative probe items; these odours were purposefully pre-selected as they were previously rated as difficult to verbalise These ratings were obtained from an earlier study that collected normative data for 200 odours (Chapter 2). The verbalisability rating used for stimulus selection was scored from 0-3 according to the quality of the verbal labels provided, with a lower score indicating vague or absent verbalisability. The 20 odours selected as negative probe items for the present study scored between 1.20 and 1.84 (M = 1.61, SD = 0.19), meaning that whilst some verbalisation may be possible, the labels tend to be vague or only refer to a broad category descriptor.

5.1.1.3 Design

A single yes/no recognition paradigm was employed to investigate the effect of PI on olfactory memory. Participants received 40 trials, where each comprised a sequence of 4 odours followed by a single test-probe. In one half of the trials, the probe was an odour presented in the preceding sequence (positive probe) and in the other half of trials the probe was not presented in the preceding sequence (negative probe). In the 20 positive probe trials (P+), each of the four serial positions were tested on five different trials (i.e. the P+ tested each of the four serial positions an equal number of times). The to-be-remembered sequence was unique, meaning that these items and the corresponding positive probes had not been experienced in any preceding trial. However, the principal independent variable concerned the 20 negative probes (P-): for 10 trials the probe was taken from the trial sequence immediately preceding the present trial (the 'recent negative' probe: RN). For the remaining 10 trials the probe was taken from the trial sequence that occurred three trials prior to the present trial (the 'non-recent negative' probe: NRN). It is these negative probes that were the hard-to-verbalise odours described above.

Three dependent variables were recorded. The yes/no response accuracy was recorded as the number of hits (correct positive probe recognition) and false- alarms (incorrect negative probe recognition). In addition, ratings were taken of the participant's confidence that their response was correct, ranging from 1 (least confident) to 5 (most confident).

5.1.1.4 Procedure

Written consent was gained from all participants prior to testing. The memory tasks were performed in a well-ventilated laboratory. Participants sat opposite the experimenter, separated by an obfuscation screen. Throughout testing, participants were instructed to focus on a fixation cross located on this screen to minimise visual interference. Participants received four blocks of ten trials, with each block separated by a 5 minute resting period in which they were able to drink water. For each trial, a sequence of four odours was presented followed by a probe odour. Each odour was presented birhinally for 2-seconds during which participants were instructed to inhale deeply. A 2 second inter-stimulus interval (ISI) separated the presentation of each odour within the sequence, followed by a 5-second retention interval prior to presentation of the test probe. Participants were required to indicate verbally (yes/no) with respect to whether the probe odour had been present in the immediately preceding sequence. Participants additionally provided a verbal confidence rating from 1 (guess) to 5 (certain). Each trial was separated by 3 seconds and the completed task lasted 37 minutes.

Figure 12 provides a schematic of the trial structure and different types of trials. The composition of the sequences and the order of trial types were pseudo-randomised with the following restrictions: (1) In order to maximise the recency of the RN probe, the probe was always taken from the third or fourth serial position of the preceding trial. Consequently, the mean number of intervening items between original presentation of the odour and re-presentation of that odour as the RN probe was 5.5 (in addition, the RN could never follow a trial where that item was also used as the test probe in a positive probe trial). An example is seen in Figure 12 where the RN probe in trial 3 ('pear') was the fourth item in the TBR sequence of trial 2. In this example, there were 5 odours intervening between original presentation of 'pear' in trial 2 and use of 'pear' as the negative probe in trial 3. (2) The NRN probe was originally presented three trials prior to the current trial, and was taken from the third or fourth serial position of that trial. Consequently, the mean number of intervening items between original presentation of the odour and re-presentation as the NRN probe was 15.5. An example is seen in Figure 12 where the NRN probe in trial 4 ('honey') was the third item in the TBR sequence of trial 1. In this example, there were 16 odours intervening between original presentation of 'honey' in trial 1 and use of 'honey' as the negative probe in trial 4. (3) Presentation of the negative probe and its original presentation in a previous trial could not overlap blocks, since the 5-minute inter-block interval would affect temporal recency of the negative probe. As a result, a block of trials could not begin with a negative probe.

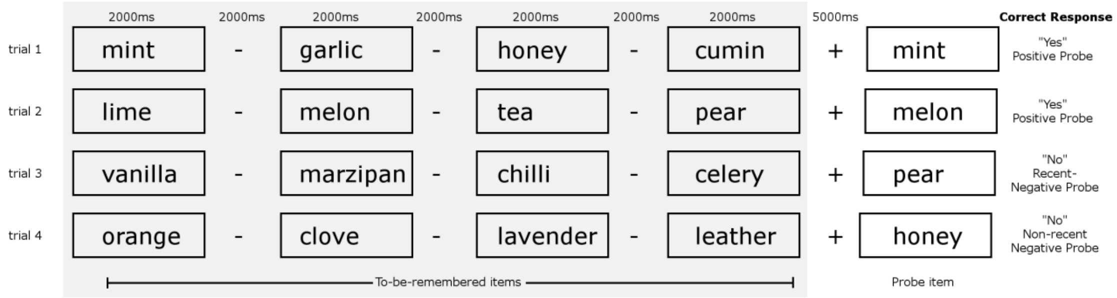

Figure 12. Schematic diagram of the recent probes task. The negative probe is taken from the immediately preceding trial in trial 3 (recent), and from three trials previously in trial 4 (non-recent).

5.1.2 Results

5.1.2.1 Recognition Sensitivity

Signal detection theory was used to determine that correct task performance exceeded chance. Overall response sensitivity (*d'*) was computed using the proportion of hits (*H*) and false alarms (*FA*), $d' = \Phi\text{-}1(H) - \Phi\text{-}1(FA)$ (Stanislaw & Todorov, 1999). Perfect scores for hits and correct rejections were adjusted by subtracting 1/(2*N*), where *N* equals the number of possible hits or correct rejections (Macmillan & Kaplan, 1985). This correction is required because *d'* is indeterminate for perfect hit or false alarm rates due to an infinite z-score (Stanislaw & Todorov, 1999). Consequently, one assumes the 'true' hit or false alarm rate is somewhere between one error and a perfect score, settling on half a miss/false alarm. A one sample t-test revealed that the d' recognition score ($M = 0.85$, $SD = 0.46$) was significantly above zero (i.e. zero = no sensitivity), $t(23) = 9.07$, $p < .001$, $d = 3.78$, meaning that participants were able to perform the task above chance. In order to assess the possibility of olfactory fatigue (and/or non-specific PI), a one-way analysis of variance (ANOVA) compared *d'* across the four experimental blocks and revealed a non-significant difference, $F(3, 69) = 0.53$, $p = .664$, $\eta_p^2 = .02$. This indicates a lack of olfactory fatigue and/or non-specific PI.

5.1.2.2 Serial position analysis

Figure 13(A) shows the serial position function for hits and reveals a recency advantage. A one-way within-participants ANOVA was conducted and revealed a main effect of serial position, $F(3, 69) = 3.83$, $p = .013$, $\eta_p^2 = .14$. Bonferroni-corrected pairwise comparisons ($\alpha = .017$) revealed a significantly greater number of hits at position 4 compared to position 2. No other comparisons were significant. In contrast to the serial position effects reported for accuracy, confidence ratings did not exhibit a main effect of serial position, $F(3, 63) = 1.13$, $p = .344$, $\eta_p^2 = .05$, and are shown in Figure 13(B).

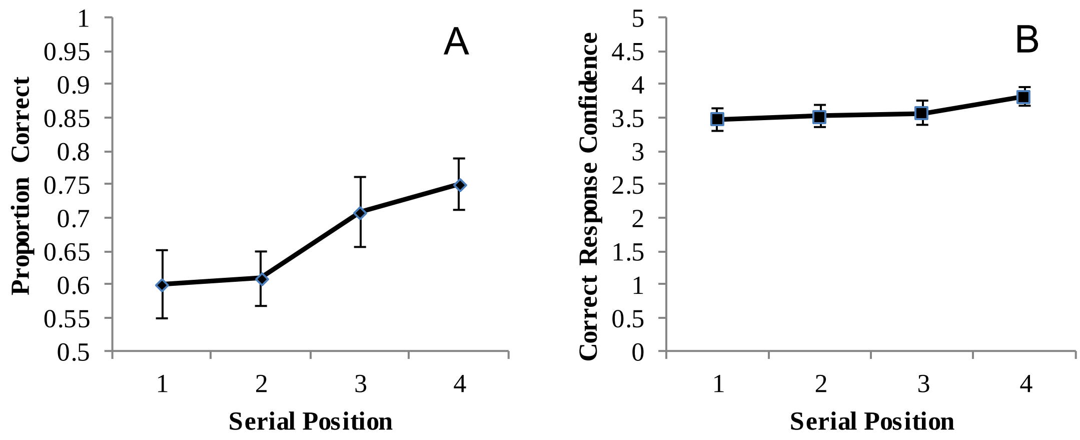

Figure 13. Serial position functions for (A) hits, and (B) confidence judgements for hits, for olfactory stimuli used in Experiment 1. Error bars denote mean standard error.

5.1.2.3 Proactive Interference

Accuracy. Table 5 displays negative probe accuracy (correct rejections) for NRN and RN probe types. A paired t-test reveals a non-significant difference between NRN (M = .65. SD = .17) and RN (M = .64, SD = .16), $t(23) = 0.09$, $p = .464$, one-tailed, $d = 0.02$, 95% CI [-.09, .10].

To further examine whether one can accept the null hypothesis of no PI, Bayes Factor (*B*) are reported, calculated using the procedures outlined in Dienes (2014). This analysis differs to *BF* reported in earlier chapters of this thesis due to the use of a plausible prior, based on previous research in the field (compared to earlier studies which used a default Cauchi prior width). Proposed cut-offs for acceptance of a hypothesis (Jeffreys, 1961), and states a *B* above 3 as providing substantial support for the alternative hypothesis, and below 1/3 provides substantial support for the null. A *B* that falls between 1/3 and 3 deems the data insensitive as to whether the alternative or null hypothesis should be accepted (where 1 equals equivalent evidence for the null and alternative hypotheses). The notation in the present study therefore follows the format $B_{H(0,\ X)}$, which refers to the specific prior used to test each hypothesis. Here, 'H' indicates a half-normal prior distribution, 'X' the predicted SD of this half-normal distribution, and the '0' signifies this comparison is against a null hypothesis of no difference.

For the present experiment, the predicted SD was 15% of the NRN probe score, taken from the non-verbal (facial stimuli) recent-probes study in Brandon et al. (2003). The SE calculated from the data was corrected to adjust for a small sample size using the formula *SE* * (1 + 20/(*df* x *df*)) (Dienes, 2008). The *B* value falls within the 'insensitive' range, $B_{H(0,\ 15\%)} = 0.46$, but does indicate that the null hypothesis (no PI) was 2.19 times more likely than the alternative.

Confidence resolution. The extent to which confidence judgments are predictive of a correct response for RN and NRN was analysed using confidence resolution. Item-level gamma correlations were calculated separately for RN and NRN probes whereby confidence rating (1-5) was correlated with accuracy (0 or 1) (see Table 5). The coefficients were then compared between the two probe types. A positive gamma coefficient reflects effective ability of confidence ratings to discriminate between correct responses and intrusion errors (Jacoby et al., 2010; Wahlheim & Jacoby, 2011). A reduction in the positive coefficient, or a negative correlation, reflects the influence of intrusion errors where a participant gives a highly confident false alarm. When calculating item-level gamma in the present and future experiments, participant data was not analysed when the responses of a participants were either all correct or all incorrect; this is because it is not possible to calculate a correlation coefficient without variance in scores. Consistent with the accuracy analysis above, there was no difference between RN and NRN in respect to confidence resolution $t(22) = 0.82$, $p = .211$, one-tailed, 95% CI [-.23, .54], $d = 0.27$. Furthermore, the Bayes Factor (using a predicted effect of 0.24 based on findings in Wahlheim & Jacoby, 2011) was insensitive, $B_{H(0,\ 0.24)} = 1.11$, indicating a lack of evidence for or against the null and alternative hypotheses.

Table 5

The proportion of correct rejections, and the confidence judgement resolution, for low verbalisability olfactory negative probes. Standard error of the mean is presented in parentheses.

	Non-Recent Negative	Recent Negative
Correct Rejections	.65 (.04)	.64 (.03)
Confidence Resolution	.20 (.12)	.05 (.12)

5.1.3 Discussion

Experiment 1 used a sequential recent probe task to examine the existence of PI in hard-to-verbalise odours. Recognition sensitivity (*d'*) was significantly above chance, demonstrating that participants were able to perform the task. Yes/no recognition performance for the 4-odour sequences revealed some evidence for recency but no primacy; a function consistent with when visual stimuli are employed in this task (Johnson, Volp, & Miles, 2014; Kerr, Avons, & Ward, 1999). However, the main focus of Experiment 1 concerned analysis of the negative probes (lures) and an absence of recent probe PI for hard-to-verbalise odours. This finding is in contrast to the prediction that hard-to-verbalise stimuli exhibit accentuated PI (see Atkins et al., 2011) and historical claims of strong PI in olfactory memory (Lawless & Engen, 1977). This is also in direct contrast to recent probe PI effects found with verbal stimuli and unfamiliar faces (e.g. Brandon et al., 2003; Craig et al., 2013; Postle et al., 2004), but is consistent with the absence of the effect with colours and shapes (Postle et al., 2004).

Importantly, whilst both the accuracy and confidence resolution analyses revealed a non-significant difference between the RN and NRN conditions, Bayes Factors showed these comparisons to be insensitive. To be clear, whilst the data clearly fails to support the existence of a difference between RN and NRN (i.e. a recent probe PI effect), the data do not provide unequivocal evidence for the null hypothesis (i.e. that RN and NRN are the same). Bayes factors revealed that for accuracy the null hypothesis was 2.19 times more likely than the alternative hypothesis, and for confidence resolution, the data supported neither the null or alternative hypothesis (B = 1.11). This suggests that there is tentative (although not strong) evidence that there is no difference between RN and NRN in respect to accuracy.

There are a number of explanations as to why PI may be absent in Experiment 1. First, it is possible that the memory task was unsuitable for olfactory stimuli, masking any PI effects. One would argue that this is unlikely because (1) performance was significantly above chance and (2) the conventional serial position function was observed for single yes/no recognition (e.g. Johnson et al., 2014; Kerr et al., 1999).

Second, the recent probe PI task typically presents the TBR memory set simultaneously at test (Jonides & Nee, 2006). Due to the constraints of olfactory perception, the memory set was presented sequentially at encoding in Experiment 1. It is possible that this methodological difference precludes PI for two possible reasons. First, in the present study there is a longer interval between initial presentation of the RN lure and its re-presentation as a test probe. Since the RN probe is less temporally recent, it may exhibit less interference. Second, with simultaneous presentation one could conceptualise the TBR items as a single item in memory, whereas for sequential presentation there are 4 discrete items presented. This would result in the RN probe in the present study being less recent in respect to the number of intervening items. The extent to which sequential presentation prevents the recent probe PI effect is addressed in Experiment 3.

A third explanation for the absence of PI in Experiment 1 concerns the characteristics of the stimuli. The negative probe odours were purposely selected for their low verbalisability score (using normative ratings from Chapter 2). The purpose of this selection was to examine PI effects for olfactory memory rather than memory for the verbal recoding of olfactory stimuli. However, it should be noted that the normative data from Chapter 2 reveal a strong positive correlation between verbalisability and familiarity (r = .88). It therefore follows that the negative probe odours used here possessed low familiarity ratings. Since unfamiliar odours are described as 'fuzzy' percepts with overlapping features (Stevenson & Mahmut, 2013a; Wilson & Stevenson, 2006), it is possible that this lack of stimulus distinctiveness may have prevented PI. As described earlier, Mecklinger et al. (2003) argues that for less distinct stimuli, the mapping in memory between the original presentation of the item and its re-presentation as the negative probe is less clear, and, as a result, the level of PI for that item is reduced. This explanation is supported by Prabhakaran and Thompson-Schill (2011) who demonstrated stronger PI effects for RN probes when using famous face stimuli (i.e. more familiar) compared to unfamiliar faces (even after accounting for verbal labelling differences).

5.2 Experiment 2: Recent-probes task with verbalisable odours

Experiment 2 is designed to examine the extent to which the absence of the PI effect in Experiment1 is due to using hard-to-verbalise odours. Using verbalisable odours for the negative probes should increase stimulus distinctiveness and accentuate PI (Mecklinger et al., 2003). Indeed, there is reason to predict differences in memory between verbalisable and hard-to-verbalise odours, since prior work has shown not only that working memory accuracy levels are higher for verbalisable odours (e.g. Jönsson, Møller, & Olsson, 2011), but that different areas of the brain area activated for these different odour types (Zelano et al., 2009). In Experiment 2, increased use of verbal codes is expected for the odours; it is therefore predicted that recent probe PI effects will be more in line with those shown for verbal stimuli (e.g. Jonides & Nee, 2006). Specifically, for verbalisable odours item-specific proactive interference effects are predicted, such that (1) there are more false positives for the RN probe compared to the NRN probe, and (2) confidence resolution is reduced for the RN probe relative to NRN probe.

5.2.1 Method

5.2.1.1 Participants

Twenty-four students from Bournemouth University (mean age = 24.96, *SD* = 6.37, females = 15, males = 9) participated in Experiment 2. As in Experiment 1, recruitment criteria required non-smokers under the age of 40 years. None had participated in

Experiment 1. The study received ethical approval via the Bournemouth University research ethics procedure.

5.2.1.2 Materials

The stimuli were as described for Experiment 1, with the crucial difference that 20 odours rated high on verbalisability were selected as the negative probe odours (see Appendix A). The verbalisability score for these odours was 0.78 *SD* above the mean for the stimulus set. Furthermore, as a manipulation check, an independent *t*-test revealed that verbalisability score for the negative probe odours used in Experiment 2 was significantly higher than the negative probe odours used in Experiment 1, $t(19) = 13.37$, $p < .001$, $d = 4.23$. In addition, the odours used for Experiments 1 and 2 also significantly differed on familiarity, $t(38) = 6.17$, $p < .001$, $d = 1.95$.

Additional efforts were made to match the odours used as RN and NRN probes on intensity, pleasantness, and irritability normative data (Chapter 2). These comparisons revealed for no differences between groups (Table 6).

Table 6

Comparison of mean normative scores (Chapter 2) for olfactory stimuli in Experiment 1 and 2 (p values are presented in parentheses).

	Experiment 1	Experiment 2	*t*-test	$B_{N(0, 1)}$
Verbalisability	1.61	2.34	-13.37 (< .001)	-
Familiarity	3.33	4.24	-6.48 (< .001)	-
Intensity	5.00	5.14	-0.55 (.585)	0.27
Pleasantness	3.49	3.68	-0.54 (.601)	0.38

Irritability	4.22	4.00	0.70 (.491)	0.38

5.2.1.3 Design

The design followed that described for Experiment 1.

5.2.1.4 Procedure

The procedure followed that described for Experiment 1.

5.2.2 Results

The same analyses were conducted as described for Experiment 1.

5.2.2.1 Recognition sensitivity

A one sample t-test revealed that the *d'* recognition score (M = 1.05, SD = 0.39) was significantly above zero (i.e. zero = no sensitivity), $t(23) = 13.24$, $p < .001$, $d = 5.52$, meaning that participants were able to perform the task above chance. A one-way analysis of variance (ANOVA) compared *d'* across the four experimental blocks and revealed a non-significant difference, $F(3, 69) = 0.53$, $p = .672$, $\eta_p^2 = .02$. This, consistent with Experiment 1, indicates a lack of olfactory fatigue and/or non-specific PI.

5.2.2.2 Serial position analysis

Figure 14(A) shows the serial position function for hits and, consistent with the findings of Experiment 1, reveals a recency advantage. A one-way within-participants ANOVA was conducted and revealed a main effect of serial position, $F(3, 69) = 7.08$, $p < .001$, $\eta_p^2 = .53$. Bonferroni-corrected pairwise comparisons ($\alpha = .017$) revealed a significantly greater number of hits at position 4 compared to positions 1 and 2. No other comparisons were significant. In contrast to the serial position effects reported for accuracy, confidence ratings did not exhibit a main effect of serial position, $F(3, 63) = 0.67$, $p = .572$, $\eta_p^2 = .03$, and is shown in Figure 14(B).

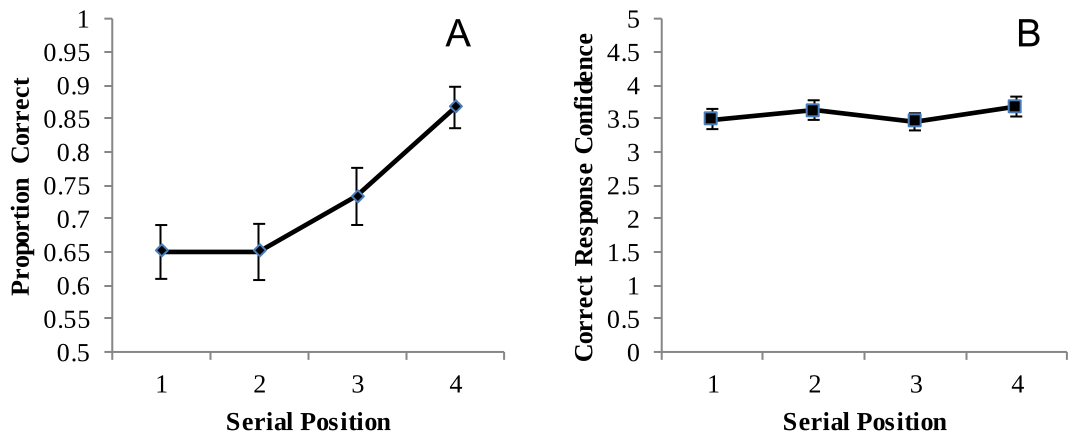

Figure 14. Serial position functions for (A) hits, and (B) confidence judgements for hits, for olfactory stimuli used in Experiment 2. Error bars denote mean standard error.

5.2.2.3 Proactive Interference

Table 7 displays the mean accuracy and confidence resolution for the NRN and RN probes. A paired *t*-test reveals borderline statistically significant lower accuracy for the RN probes ($M = .60$, $SD = .19$), compared to NRN probes ($M = .70$, $SD = .14$), $t(23) = 2.037$, $p = .027$, one-tailed, $d = 0.85$, 95% CI [.00, .18].

This effect is further supported by the Bayes factor which shows that the alternative hypothesis is 3.96 times greater than the likelihood of the null ($B_{H(0,\ 15\%)} = 3.96$). These findings indicate a PI effect for recognition accuracy.

Consistent with the accuracy data, confidence resolution was significantly lower for RN probes ($M = -.11$, $SD = .56$), compared to that found with NRN probes ($M = .33$, $SD = .33$), $t(21) = 3.63$, $p = .001$, one-tailed, $d = 0.95$, 95% CI [0.19, 0.69], $B_{H(0,\ 0.24)} = 103.53$. Taken together, this is strong evidence for a recent probe PI effect with verbalisable odours.

Table 7

The proportion of correct rejections, and the confidence judgement resolution, for high verbalisability olfactory negative probes. Standard error of the mean is presented in parentheses.

	Non-Recent Negative	Recent Negative
Correct Rejections	.70 (.03)	.60 (.05)
Confidence Resolution	.33 (.07)	-.11 (.12)

5.2.3 Discussion

Experiment 2 reports recent probe PI effects for verbalisable odours both in respect to the correct rejection of negative probes and in respect to confidence resolution. That is, false alarms were greater for recent (RN) lures compared to NRN lures, and the positive relationship between confidence and accuracy seen for NRN lures was significantly lower, and negative, for RN lures. This is in stark contrast to Experiment 1, where, for hard-to-verbalise odours, evidence was weighted towards the null hypothesis of no PI effects. In discussing these findings one must firstly consider why hard-to-verbalise and verbalisable odours differ in PI, and secondly, explore how this speaks to recent probe PI effects with other stimulus types.

First, it is intriguing that verbalisable and hard-to-verbalise odours exhibited different PI effects given similarities in performance across studies. That is, (1) participants performed in both studies above chance, (2) recognition sensitivity did not differ significantly between both tasks ($t(46) = 1.63$, $p = .110$, $d = 0.47$), though the data were insensitive to differences ($B_{N(0,\ 0.3)} = 1.18$), and (3) the serial position functions for hits were qualitatively equivalent (recency but no primacy). As outlined earlier, it is possible that stimulus characteristic differences in respect to the distinctiveness of verbalisable and hard-to-verbalise odours, underpinned the difference (Mecklinger et al., 2003). It should be noted that the hard-to-verbalise odours are also less familiar (Chapter 2), and representations for less familiar odours are argued to be less distinct with overlapping features (Stevenson & Mahmut, 2013a; Wilson & Stevenson, 2006). If the negative probe odours are less defined/distinct, it is possible that participants are less aware/confident that the current probe odour maps onto the representation for that same odour in the preceding trial (the RN probe) (an account outlined by Mecklinger et al., 2003). The present findings may therefore not be attributed to stimulus verbalisability specifically, but from a reduction in stimulus distinctiveness from some combination of familiarity and verbalisability differences.

Second, whilst the present findings fit with the Mecklinger et al. explanation, they are not necessarily consistent with recent probe PI effects found with other stimulus effects. For example, Atkins et al. (2011) showed greater PI effects during an articulatory suppression task, though they do suggest this is due to reduced distinctiveness of the episodic information rather than the verbalisability of the stimulus. However, Postle et al. (2004) reported no behavioural evidence for PI with shapes and colours; for which one might expect colours in particular to have easily accessible verbal labels. In addition, PI effects have been found with faces regardless of verbalisation (Brandon et al., 2003); a stimulus that is argued to be processed similarly to that of odours (Kärnekull et al., 2015).

Any attempts at cross-modal comparisons with olfactory PI effects are, however, confounded by methodological differences. To reiterate, the recent probe PI procedure involves sequential presentation of the TBR memory set. In contrast, the classic version of the task involves simultaneous presentation of the memory set (see Craig et al., 2013; Jonides & Nee, 2006). It is therefore possible that differences in the encoding experience of the memory set affects the magnitude of PI. To be clear, whilst Experiment 1 suggests that hard-to-verbalise odours differ from, for example, words (Jonides & Nee, 2006) and faces (Brandon et al., 2003), in not demonstrating a recent probe PI effect, it is possible that this apparent difference is underpinned by differences in method rather than stimulus. The sequential presentation method employed in Experiments 1 and 2 has two important implications. First, the time between initial odour presentation and the recent-negative probe item is greater than those typically seen in the recent-probes task (e.g. Badre & Wagner, 2005; Craig et al., 2013). Second, the relative isolation of the TBR items may have had some unknown effect on the item-specific interference (though it should be noted that Monsell, 1978, did use sequential presentation of verbal stimuli in their seminal work). As a consequence, other stimulus types are applied to the sequential recent probe task in order to make meaningful comparisons with the olfactory PI findings reported in Experiments 1 and 2.

5.3 Experiment 3: Recent-probes task with verbal stimuli

Experiments 3 and 4 apply visual-verbal (words) and visual non-verbal (faces) stimuli to the sequential recent probes tasks used in Experiments 1 and 2. These stimulus types have been shown to exhibit recent probe PI effects when the memory set are presented simultaneously (Brandon et al., 2003; Craig et al., 2013; Postle et al., 2004). If these stimulus types also show PI for the sequential version of the task, it will demonstrate that hard-to-verbalise odours (Experiment 1) differ to other stimulus types in respect to the presence of PI.

Experiment 3 used words[3] and based upon previous work showing PI effects with verbal stimuli (e.g. Jonides & Nee, 2006), PI was predicted to be evidenced by both higher false alarms and lower confidence resolution for the RN probe relative to the NRN probe.

5.3.1 Method

5.3.1.1 Participants

Twenty-four students from Bournemouth University (mean age = 23.58, *SD* = 9.55, females = 20, males = 4) participated. None had participated in Experiments 1 or 2. The study received ethical approval via the Bournemouth University research ethics procedure.

---

[3] Pilot testing using the presentation times employed in Experiments 1 and 2 (2 s exposure time and a 2 s ISI) revealed that direct methodological replication using verbal stimuli is unsuitable due to ceiling effects. Ceiling effects were also produced in two further pilots in which (1) presentation times for each stimulus item were reduced with ISIs increased (thereby maintaining the same temporal interval between re-presentations of the negative probe used in Experiments 1 and 2) and (2) reduced presentation times and ISIs, with increased inter-trial interval increased (again maintaining the same temporal interval between re-presentations of the negative probe used in Experiments 1 and 2). Consequently, presentation times, ISI, and time between trials in Experiment 3 were all proportionately reduced to increase task difficulty (i.e. presentation-interval ratios were identical to Experiments 1 and 2; however absolute timings were reduced). This resulted in a reduction in the time elapsed between initial presentation and re-presentation as a negative probe (a mean of 3.45s between TBR item and RN probe in the present experiment, compared to a mean of 31s for Experiments 1 and 2), meaning the large temporal distance between items as a potential confound was not addressed in this experiment (to pre-empt, this issue is addressed in Experiment 4). However, the number of items between initial presentation of the item and re-presentation as the negative probe remained identical to that described for Experiments 1 and 2.

5.3.1.2 Materials

Verbal stimuli were 656 high frequency, concrete nouns, selected from the N-Watch default vocabulary of 30,605 words (Davis, 2005). Mean word length was 6.4 letters (min = 6, max = 7), with 2 syllables. Minimum CELEX frequency (per million words) was 1.62 with a mean of 39.12 (*SD* = 88.06). These words were presented in the centre of a 22 inch monitor in size 18pt. The open-source experimental presentation software OpenSesame was used to present words and record responses.

5.3.1.3 Design

The design was as described for Experiments 1 and 2.

5.3.1.4 Procedure

The recent-probes procedure from Experiment 1 and 2 was adapted for verbal stimuli. Due to the brevity of the task, a total of 160 trials were presented to participants, with each trial comprising 4 TBR items followed by a single yes/no recognition probe.

Testing took place at Bournemouth University in an individual laboratory booth. Participants gave written consent, and were instructed on the task procedure. A short (15 trial) practice task preceded the testing phase in order to familiarise participants with the speeded presentation of items. Each TBR item was presented for 100ms, with a 100ms ISI presented as a fixation cross. A 250ms fixation interval separated the final TBR item and presentation of the recognition probe item. When this probe item appeared on the screen, participants were required to make a key press of ‘Z’ to make a negative response, and a ‘V’ key press for a positive response. Both decisions were made with the left hand. Immediately following the set membership decision, the probe was removed and ‘Confidence?’ appeared centrally on the monitor, prompting a numerical confidence rating of 1 (guessing) to 5 (certain) to be made with the right hand on the number keypad. Participants were advised to make their responses to both decisions as quickly and accurately as possible. A 150ms fixation cross separated trials following both responses, and an enforced 30 second break was included every 25 trials. The testing lasted approximately 15 minutes.

5.3.2 Results

To enable direct comparison with Experiments 1 and 2, only the first 20 positive and negative probes (10 RN and 10 NRN) were analysed.

5.3.2.1 Recognition sensitivity

A one sample t-test revealed that the *d'* recognition score (M = 2.35, SD = 0.71) was significantly above zero (i.e. zero = no sensitivity), $t(23) = 16.26$, $p < .001$, $d = 6.78$, meaning that participants were able to perform the task above chance. Item-nonspecific PI was analysed by splitting responses into 10-trial blocks (with the first 5 negative and positive probes allocated to block 1, etc.) and calculating *d'*. A one-way analysis of variance (ANOVA) compared *d'* across the four experimental blocks and revealed a non-significant difference, $F(3, 69) = 1.04$, $p = .380$, $\eta_p^2 = .04$. This indicates a lack of task fatigue and/or non-specific PI.

5.3.2.2 Serial position analysis

Figure 15(A) shows the serial position function for hits and, consistent with the findings of Experiments 1 and 2, reveals recency but no primacy. A one-way within-participants ANOVA was conducted and revealed a main effect of serial position, $F(3, 69) = 4.64$, $p = .005$, $\eta_p^2 = .17$. Bonferroni-corrected pairwise comparisons ($\alpha = .017$) revealed a significantly greater number of hits at position 4 compared to positions 1 and 2. No other comparisons were significant. As reported for Experiments 1 and 2, serial position effects for confidence ratings were non-significant, $F(3, 63) = 2.01$, $p = .121$, $\eta_p^2 = .08$, shown in Figure 15(B).

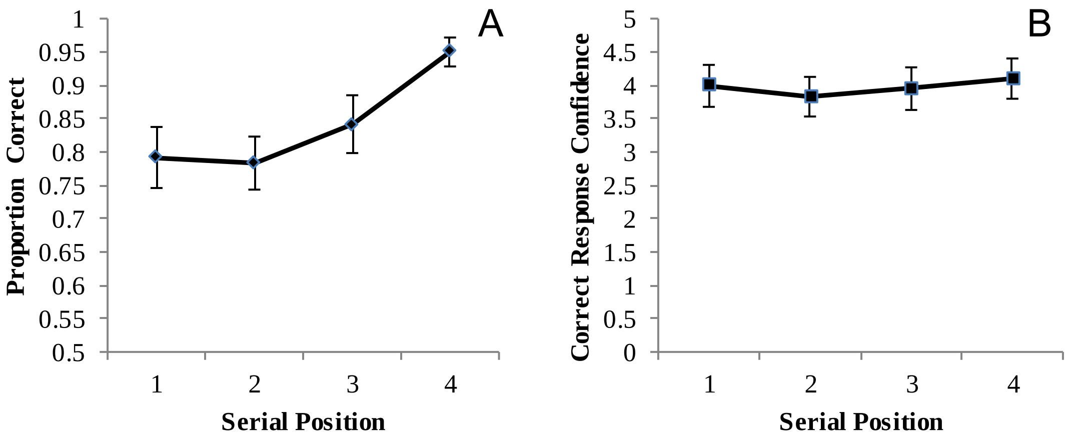

Figure 15. Serial position functions for (A) hits, and (B) confidence judgements for hits, for verbal stimuli used in Experiment 3. Error bars denote mean standard error.

5.3.2.3 *Proactive Interference*

Table 8 displays the mean accuracy and confidence resolution for the NRN and RN probes. A paired t-test reveals statistically significant lower accuracy for the RN probes (M = .84, SD = .13), compared to NRN probes (M = .90, SD = .12), $t(23) = 2.89$, p = .004, one-tailed, d = 0.54, 95% CI [.02, .11]. Bayes factor was again computed using a predicted drop of 10% for the RN probes; this is based upon verbal recent-probe findings (Craig et al., 2013; Jonides & Nee, 2006; Monsell, 1978). The Bayes factor provided strong support for the alternative hypothesis ($B_{H(0,\ 7\%)} = 21.06$). These findings indicate a strong PI effect for recognition accuracy.

In contrast to the accuracy data, confidence resolution did not significantly differ between the RN (M = .14, SD = .57) and NRN probes, (M = -0.33, SD = 0.73), $t(10)$ = 1.76, p = .055, one-tailed, d = 0.72, 95% CI [-0.13, 1.07], $B_{H(0,\ 0.7\%)} = 1.91$.

Table 8

The proportion of correct rejections, and the confidence judgement resolution, for verbal stimuli negative probes. Standard error of the mean is presented in parentheses.

	Non-Recent Negative	Recent Negative
Correct Rejections	.90 (.02)	.84 (.03)
Confidence Resolution	.14 (.17)	-.33 (.22)

5.3.3 Discussion

Experiment 3 provides strong support for recent probe PI effects for words using the sequential presentation method employed in Experiments 1 and 2. This strong effect was found in respect to accuracy. It is, however, also worth noting that PI effects were not found for words in respect to confidence resolution. This could be explained by low statistical power for this analysis. That is, calculating confidence resolution necessitates that the participant provides both correct and incorrect responses for the RN and NRN probes. However, since 13 participants provided perfect performance for the RN and/or the NRN probes, this prevented confidence resolution from being calculated for those participants. As a result, the statistics are based upon a sample of 9. Nevertheless, it should be noted that for the analysed data, the effect size was large (d = 0.72), providing some support for a PI effect with confidence resolution.

The findings of Experiment 3, therefore, demonstrate that the apparent difference in susceptibility to PI for odours (Experiment 1) and verbal stimuli (Jonides & Nee, 2006) cannot be explained by the sequential presentation of the TBR items. The experiment has shown that the effect remains for verbal stimuli when the TBR items are presented sequentially rather than simultaneously. It is, however, of interest to note that whilst PI was absent for Experiment 1 but observed for words in the present study, both hard-to-verbalise odours and words exhibited qualitatively equivalent serial position functions (recency but no primacy: a function consistent with previous single yes/no recognition studies with verbal stimuli, Brian McElree & Dosher, 1989; Monsell, 1978). This shows some similarity in the memory functioning of hard-to-verbalise odours and words consistent with other recognition tasks (e.g. A. J. Johnson & Miles, 2007).

Notwithstanding the use of sequential presentation in Experiments 1 and 3, there remains an important methodological difference between the two experiments. To avoid ceiling effects, the words were presented at a faster rate (100ms with a 100ms ISI) than the odours (2s with a 2s ISI). As a result, the average interval between initial presentation of an item and its re-presentation as the NRN or RN probe differed dramatically for words (NRN interval = 9.65s, RN interval = 3.45s) compared to odours (NRN interval = 79s, RN interval = 31s). If one assumes that PI reduces over time (an assumption on which the recent probe task is premised), then Experiment 1 is weighted against observing a PI effect for hard-to-verbalise odours relative to words in Experiment 3. Consequently, the difference between hard-to-verbalise odours and words in respect to PI may be due to presentation intervals rather than stimuli per se.

5.4 Experiment 4: Recent-probes task with face stimuli

Experiment 4 addresses the criticism that PI effects were not found for hard-to-verbalise odours (Experiment 1) due to long presentation intervals between initial presentation of the item and its re-presentation as the negative probe. This experiment employs faces as TBR stimuli for three reasons: (1) faces have previously exhibited a recent probe PI effect (Brandon et al., 2003), (2) pilot work revealed performance not to be at the ceiling when timings more closely reflected that used in Experiments 1 and 2, and (3) faces are often considered a suitable comparison stimuli for odours due to possible holistic processing for both classes of stimuli (Kärnekull et al., 2015; Stevenson & Mahmut, 2013b). To ensure that the presentation intervals matched that described in Experiments 1 and 2, a 24s interval was introduced between each trial. As in the previous experiments, PI is examined by comparison between the RN and NRN probes in respect to both accuracy and confidence resolution.

5.4.1 Method

5.4.1.1 Participants

Twenty-four students from Bournemouth University (mean age = 20.76, *SD* = 4.88, females = 22, males = 2) participated. None had participated in Experiments 1-3. The study received ethical approval via the Bournemouth University research ethics procedure.

5.4.1.2 Materials

One-hundred and sixty faces were randomly selected from the Glasgow Unfamiliar Face Database (Burton, White, & McNeill, 2010). All .jpg images were 350px x 473px, showed a front (full face) view with a neutral expression, and were cropped to remove any visible clothing. Stimuli were presented in the centre of a 22-inch 60Hz monitor and responses collected using OpenSesame (Mathôt et al., 2012).

5.4.1.3 Design

The same design was used as described for Experiments 1-3.

5.4.1.4 Procedure

Participants were tested individually in a laboratory booth at Bournemouth University. Participants were presented with on-screen instructions, and initiated the task by pressing the space bar. The trial started 1s after initiating the task, where four TBR items were displayed sequentially for 250ms, with an ISI of 250ms. A retention interval of 625ms separated the final TBR item and presentation of the recognition probe. The response format followed that described for Experiment 3. Following these responses, a 24s delay (required to approximately match the total time of 29 seconds between initial presentation in the fourth serial position and RN probe, in Experiments 1 and 2) separated trials. Participants were warned 1 second before the beginning of the next trial by a change in the number of fixation dots in the centre of the screen. A total of 40 trials were performed, matching the 20 positive and 20 negative probes presented in Experiments 1 and 2. A 1-minute break separated each block of 10 trials.

5.4.2 Results

5.4.2.1 Recognition sensitivity

A one sample t-test revealed that the *d'* recognition score (*M* = 2.22, *SD* = 0.45) was significantly above zero (i.e. zero = no sensitivity), $t(23) = 24.19$, $p < .001$, $d = 10.09$. A one-way analysis of variance (ANOVA) compared *d'* across the four experimental blocks and revealed a non-significant difference, $F(3, 69) = 0.91$, $p = .441$, $\eta_p^2 = .04$. This indicates a lack of task fatigue and/or item non-specific PI.

5.4.2.2 Serial position analysis

Figure 16(A) shows the serial position function for hits and, consistent with the findings of Experiments 1-3, reveals recency but no primacy. A one-way within-participants ANOVA was conducted and revealed a main effect of serial position, $F(3, 69) = 3.37$, $p = .023$, $\eta_p^2 = .13$. Bonferroni-corrected pairwise comparisons ($\alpha = .017$) revealed a significantly greater number of hits at position 4 compared to position 2. No other comparisons were significant. Shown in Figure 16(B), confidence ratings also revealed a significant main effect of serial position, $F(3, 63) = 13.17$, $p < .001$, $\eta_p^2 = .36$, with a strong recency effect evidence (position 4 > positions 1-3).

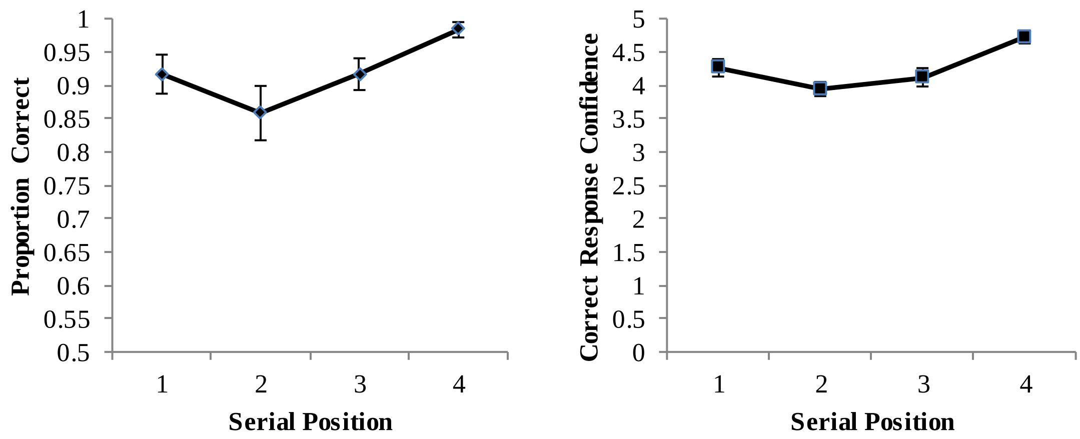

Figure 16. Serial position functions for (A) hits, and (B) confidence judgements for hits, for facial stimuli used in Experiment 4. Error bars denote mean standard error.

5.4.2.3 Proactive Interference

Table 9 displays the mean accuracy and confidence resolution for the NRN and RN probes. A paired t-test reveals statistically significant lower accuracy for the RN probes ($M = .70$, $SD = .16$), compared to NRN probes ($M = .80$, $SD = .16$), $t(23) = 3.09$, $p = .003$, one-tailed, $d = 0.63$, 95% CI [.03, .16]. Bayes factor was again computed using a predicted drop of 15% for the RN probes (based on the recent probe PI effect reported for faces, Brandon et al., 2003). The Bayes factor provided strong support for the alternative hypothesis ($B_{H(0,\ 15\%)} = 32.43$). These findings indicate a strong PI effect for recognition accuracy.

In contrast to the accuracy data, confidence resolution did not significantly differ between the RN ($M = .03$, $SD = .42$) and NRN probes, ($M = 0.13$, $SD = 0.56$), $t(18) = 0.67$, $p = .256$, one-tailed, $d = 0.22$, 95% CI [-0.23, 0.44], $B_{H(0, 0.24)} = 0.92$.

Table 9

The proportion of correct rejections, and the confidence judgement resolution, for face stimuli negative probes. Standard error of the mean is presented in parentheses.

	Non-Recent Negative	Recent Negative
Correct Rejections	.80 (.03)	.70 (.03)
Confidence Resolution	.13 (.13)	.03 (.10)

5.4.3 Discussion

Experiment 4 provides a demonstration of recent probe PI with non-verbal visual stimuli (faces) using the sequential presentation method of the TBR items. This demonstrates that the recent probe PI effect found with faces (Brandon et al., 2003) is not reliant on simultaneous presentation at encoding. In addition, the PI effect was found despite lengthened intervals between presentation of an item and its re-presentation as a negative probe. A 24s interval was introduced between the trials to ensure parity with Experiment 1 in respect to the timings of each trial. Following this manipulation one can conclude that, in contrast to faces, hard-to-verbalise odours exhibit no recent probe PI effect, and that this difference in neither a result of sequential presentation of the TBR items nor due to lengthened intervals between the re-presentation of items.

It should, however, be noted that PI was not reported for confidence resolution. Unlike Experiment 3, this non-significant effect does not appear to be due to reduced statistical power. Indeed, the effect size (d = 0.22) for the confidence interval resolution was notably smaller than Experiments 2 and 3 (d = 0.95 and 0.72, respectively). It is unclear why in Experiment 4, accuracy but not confidence resolution supports PI.

5.5 General Discussion

This chapter has described four experiments which have examined cross-modal PI effects using the recent probes task. Evidence for PI was found with verbalisable odours (Experiment 2), words (Experiment 3), and faces (Experiment 4), but no PI for hard-to-verbalise odours (Experiment 1). Such a finding may provide support against a domain-general effect in the recent-probes task (Jonides & Nee, 2006; Leung & Zhang, 2004). In addition, the findings with words and faces extend the established recent probes effect (Brandon et al., 2003; Craig et al., 2013; Jonides & Nee, 2006) to a version of the task in which the TBR memory set is presented sequentially (Experiments 3 and 4).

It is unclear why low verbalisability odours have not demonstrated recent probe PI effects. However, the contrasting PI findings of Experiments 1 (low verbalisability odours) and 2 (verbalisable odours) are consistent with other studies showing memory differences for verbalisable and hard-to-verbalise odours (e.g. Jönsson et al., 2011; Zelano et al., 2009). A prosaic explanation for the current data is that olfactory stimuli do not elicit item-specific PI effects, and that the PI effects in Experiment 2 are illustrative of verbal memory following (at least partial) verbal recoding of the odours. The different effects may therefore be due to stimulus characteristics that relate to the level of stimulus verbalisability, and perhaps related to whether a stored olfactory representation is consciously accessible in odour memory (see Chapters 3 and 4). However, the extent to which odours in the current study are verbalisable correlates strongly with familiarity (r = .88; Chapter 2). As noted earlier, it is possible that the reduced distinctiveness of these unfamiliar odours resulted in disrupted matching in memory between the original presentation of the item and its re-presentation as the negative probe (Mecklinger et al., 2003). This reduced matching ability would attenuate PI effects.

It should be noted that the present findings contradict that of Köster et al. (2002) who examined PI for odours using implicit memory. In their study, PI for the incidental associations between odours and rooms was only found when the odours were not identified. In contrast, the present study only reported PI for verbalisable odours. It is difficult to make such cross-study comparisons, but these findings may suggest that the verbalisability of the odour has different effects for explicit and implicit memory tasks. Moreover, it is suggested by Köster et al. that such explicit memory tasks may be unsuitable for odours since the experience and learning of odours in everyday life is epiphenomenal (i.e. the present odour is incidentally associated with the present experience). Consequently, when odours are tested explicitly it encourages participants to employ verbal coding, with the task becoming a de facto measure of verbal memory. However, one would argue that this issue with explicit odour memory tasks has been mitigated through the employment of hard-to-verbalise odours in Experiment 1 (odours previously identified as hard-to-verbalise in a large scale normative study, Chapter 2). Indeed, whilst it remains possible that participants attempt to employ rudimentary labels for these odours, the differences between Experiment 1 and 2 in respect to the presence of PI may suggest that: (1) hard-to-verbalise and verbalisable odours are affected differently by PI, and (2) verbalisable odours may well be represented in part by a verbal or semantic code.

If one accepts that (non-verbal) olfactory stimuli are not susceptible to item-specific PI, it suggests qualitative differences between olfactory memory and other stimulus types. This difference is curious given similarities in the serial position functions produced across stimuli in the present study for single yes/no recognition (recency but not primacy, consistent with Hay, Smyth, Hitch, & Horton, 2007; Johnson et al., 2014; Kerr et al., 1999). However, beyond the present study, memory for olfactory stimuli has produced serial position functions that appear qualitatively different to other stimulus types (Johnson et al., 2013; Johnson & Miles, 2009; Reed, 2000). These qualitative differences for olfactory memory add credence to the proposal that olfactory stimuli are represented within a modality specific store (Andrade & Donaldson, 2007), but that verbal dual-coding of the stimuli can supplement the perceptual representation (Paivio, 1990; Yeshurun, Dudai, & Sobel, 2008).

The present study demonstrates that the characteristics of the olfactory stimuli can determine the reported memory effects (a claim also made for other stimulus types, e.g. Hay et al., 2007; Horton, Hay, & Smyth, 2008; Rajaram, 1998). This finding adds weight to the proposal in Chapter 2 that any attempts at investigating whether olfactory memory operates analogously to other stimulus types must consider the characteristics of the odours used in that study.

5.6 Conclusion

In summary, the present study demonstrates item-specific PI effects in verbal memory, visual (face) memory, and in memory for familiar (verbalisable) olfactory stimuli. However, these effects are not found with hard-to-verbalise odours. The study therefore adds weight to the proposition that differences exist in terms of how verbalisable and hard-to-verbalise odours are represented (Zelano et al., 2009) and suggest that making comparisons between olfactory memory and that of other stimulus types might be affected by the choice of odours; a control that requires further consideration in future work.

Chapter 6: General Discussion

6 Chapter summary

This chapter begins with a review of the thesis objectives, and is followed by a summary and discussion of the findings in relation to views of olfactory memory.

6.1 Research aims

The thesis had two main overarching aims. The first aim was to obtain normative data on a large set of odours. There are conflicting findings in the olfactory memory literature regarding qualitative differences between odour memory and other modalities, and it is possible that these differences have arisen from poor control over stimulus characteristics. The second aim of the thesis was to examine ability for a perceptually-based olfactory working memory, comparable to working memory for verbal and visual stimuli. Using the normative data obtained, odours were manipulated on dimensions related to verbalisability in order to investigate representational changes, strategy adoption, and the application of controlled working memory resources. These aims are outlined in more detail below.

6.1.1 Normative data: Individual differences in olfactory perceptual experience

Individual differences are undoubtedly important when examining odour representations in memory (Kaeppler & Mueller, 2013), and is particularly relevant due to the proposed influence of prior experience and top-down processes on perception (Wilson & Stevenson, 2006). Studies of memory and perceptual processing in non-olfactory modalities have greatly benefitted from normative data that allow the control and manipulation of pertinent dimensions (e.g. Coltheart, 1981; Yoon et al., 2004).

However, in olfaction, the creation of such databases is typically limited to a small number of dimensions and odorants (Sulmont, Issanchou, & Koster, 2002), or to a limited category of odours (e.g. familiar odours, Doty, Shaman, Kimmelman, & Dann, 1984). An aim of this thesis was therefore to create a large database of commercially available odours, and to assess the utility of these normative data for odour classification in memory experiments.

6.1.2 Representation of odours in working memory

Using the normative data, an aim of the study was to investigate how odours are represented in working memory. Jönsson et al. (2011) demonstrated a working memory advantage for verbalisable odours, but nevertheless reported above chance performance for hard-to-name odours. This finding suggests that although explicit olfactory memory is difficult to disentangle from verbal processes due to a tendency to label identified odours (Jönsson et al., 2011; White et al., 2015), odours can be represented perceptually in memory. Indeed, there is clear evidence for odour representations in memory that are not solely reliant on verbal or semantic information (Andrade & Donaldson, 2007; Jönsson et al., 2011; Møller et al., 2004; White et al., 1998). For example, olfactory short-term memory is unaffected by concurrent verbal tasks, supporting an ability to maintain odours without verbal mediation (Andrade & Donaldson, 2007; cf. Annett & Leslie, 1996). Furthermore, brain regions associated to verbal and olfactory processes display dissociated activation according to whether an odour is identified (Zelano et al., 2009). This is proposed to support a perceptually-based representation in olfactory short-term memory, perhaps stored in a separate olfactory buffer.

Notwithstanding evidence that supports an independent store for olfactory information, there is clearly some effect on odour memory when verbal or semantic information is available (e.g. Kärnekull, Jönsson, Willander, Sikström, & Larsson, 2015; Lyman &

McDaniel, 1986, 1990). However, evidence is equivocal regarding the importance of these additional verbal/semantic processes. One possible explanation is that a switch to a verbal-only representation occurs when verbal information is available (Herz & Engen, 1996), though this has been criticised (see White et al., 1998). Alternatively, there is evidence that verbal information forms at least some part of odour representations in memory (e.g. White et al., 1998; Yeshurun, Dudai, & Sobel, 2008; Zelano, Montag, Khan, & Sobel, 2009), and may suggest a perceptually-based odour memory that can be facilitated by an additional verbal *dual-code* (Paivio, 1990). It has also been suggested in the object-processing account of olfactory perception (Wilson & Stevenson, 2006) that odours are perceived and subsequently remembered by activating a stored, perceptually-based olfactory representation. Short-term memory, they suggest, is supported by residual activation of these stored odour patterns, but that this recognition memory may also be facilitated by activation of linked verbal or semantic information similar to the use of multiple memory traces described in dual-code theory (Paivio, 1990).

An aim of the present studies was therefore to examine evidence for differences in how odours are represented in working memory when odours have varying levels of semantic or verbal information available. Increased availability of this information has previously shown improved working memory performance (Jönsson et al., 2011), but the underlying processes that drive this improvement is unclear. This thesis examined, amongst other explanations, whether this improvement was because of a strong verbal code allowing verbal rehearsal processes to be engaged. This aim was closely tied to the examination of processes engaged in olfactory working memory, described in the section below.

6.1.3 Verbal/semantic processing effects on working memory

A key issue for working memory in the olfactory domain is that working memory is linked to the ability to rehearse items and to form an internal representation, or mental image (Stevenson, 2009; Tong, 2013). Indeed, working memory is defined by the ability to actively maintain and manipulate a stored representation (e.g. Baddeley & Hitch, 1974), but there are suggestions that olfactory memory may not engage these processes in the same way as other modalities (Stevenson, 2009). Whilst there is evidence that olfactory imagery is possible (Bensafi et al., 2007; Djordjevic et al., 2005; Rolls et al., 2008), this is thought to be strongly dependent on expertise (Arshamian & Larsson, 2014; Delon-Martin et al., 2013; Plailly et al., 2012), and it has been argued that such imagery is not consciously accessible (Stevenson, 2009; Stevenson & Attuquayefio, 2013; Zucco, 2003).

An aim of this thesis was therefore to examine the engagement of controlled working memory resources on odour memories. The non-perceptual benefits from additional information in an odour representation may have distinct effects on processing in working memory. For example, as mentioned in Section 6.1.2, the working memory advantage for verbalisable odours may follow the adoption of a verbal rehearsal strategy to maintain and update odours in memory. For odours where verbal information was weak, or unavailable, this thesis explored whether other maintenance processes were engaged. For example, an attentional refreshing process was proposed for non-verbal stimuli to maintain and update information in working memory tasks (e.g. M. R. Johnson et al., 2015). However, such a strategy is dependent on a consciously accessible internal representation (see above).

6.1.4 Automatic processing in olfactory working memory

A possible complication to understanding previous demonstrations of olfactory working memory capacity, particularly in the n-back task, is the contribution of both recollection (controlled retrieval of contextual information) and familiarity (a strength signal that may reflect automatic processes) (Loaiza, Rhodes, Camos, & McCabe, 2015; Oberauer, 2005). Consequently, although above chance performance has been observed for odours using the n-back task (Dade et al. 2001; Jönsson et al. 2011), performance may be dependent on a processes that does not reflect the controlled retrieval of items and allocation of working memory resources (Juvina & Taatgen, 2007; Kane et al., 2007). To be clear, the previously observed above-chance working memory performance for low verbalisability odours (Jönsson et al., 2011) may not have been due to working memory processes (such as the rehearsal and updating of items and their bound serial positions, Oberauer, 2005), but a familiarity strength judgment. This thesis therefore explored the contribution of automatic processing on measurements of olfactory working memory.

In Wilson and Stevenson (2006), odour recognition is proposed to be driven by residual activation of an odour object. This is effectively a familiarity-based process, where the activation strength of an item is compared to a criterion of activation to allow a decision of previous experience. This is also described as *recognition without identification* (Cleary, 2010), where a recognition judgement is made from the familiarisation of perceptual features of an odours that are reinstated upon presentation of the probe item.

Finally, episodic memory research has shown that recollection in odour memory is similar to verbal memory when items are identified (M. J. Olsson et al., 2009). Olsson et al. suggested that this is due to the conceptual salience of the representation improving recollection ability (Rajaram, 1998, see also Hay et al. 2007 for an amodal account of memory that describes the psychological distinctiveness of items). In working memory, effects of semantic information on recollection would therefore reflect the increased use of controlled working memory resources in the task (Loaiza et al., 2015). The present thesis examined the contribution of recollection and familiarity in olfactory working memory through employment of the remember-know procedure. Moreover, Chapter 3 explored whether the contribution of recollection and familiarity changes dependent on the characteristics of the odour.

6.1.5 Proactive interference as a function of verbalisability

The differences between odours low and high on verbalisability were also examined in the present thesis through their ability to elicit proactive interference. The n-back task is characterised by sequential presentation of stimuli, requiring a decision upon presentation of each item. Furthermore, the resolution of conflict between familiarity and recollection (i.e. a form of proactive interference) is an important part of any working memory strategy in this task (Kane et al., 2007). Consequently, changes in the effects of PI with differing odour representations may contribute to an explanation for differences in levels of working memory performance. Previous findings of a flat forgetting curve for odours have been attributed to strong proactive interference (PI) and weak or absent retroactive interference (Lawless & Engen, 1977). However, there has been little examination of PI for pure odour memory, compared to that explored for other modalities (but see, Köster, Degel, & Piper, 2002, for implicit memory; and see Valentin, Dacremont, & Cayeux, 2011, for non-specific PI).

This thesis applied a recent-probes paradigm to low verbalisability odour memory in order to investigate item-specific PI (see Jonides & Nee, 2006), and to compare these item-specific PI effects to observations in other modalities. Specifically, PI effects for low verbalisability odours were compared to experiments using high verbalisability odours, visual stimuli (faces), and verbal stimuli (words).

6.2 Summary of findings and implications for theory

6.2.1 Utility of normative data and the relationship between dimensions

This thesis has provided normative data for a large number of commercially available odours that can be used in experimental studies, and satisfies a clear gap in the olfactory literature that can benefit future researchers.

An important requirement for odour normative data was that the variance in ratings was attributed to differences between the odours, rather than variability due to individual differences. That is, the agreement of scores for a particular odour must be higher than the agreement of scores across all odours. Without such, any normative characteristics of the odours will be masked by individual differences. Assessment of these relationships revealed several differences across dimensions, suggesting that some dimensions may be more useful than others. Specifically, the familiarity, intensity, pleasantness, and irritability scores showed greater agreement across participants than variability across odours, suggesting normative scores based on these dimensions are suitable for controlling olfactory stimuli. In contrast, age of acquisition and complexity dimensions showed greater agreement between odours than the variability across participants, suggesting a greater effect of individual differences.

It should also be noted that in the present data there were strong correlations between dimensions deemed suitable for use as normative data (e.g. familiarity), and dimensions that saw a large effect of individual differences (e.g. verbalisability). This suggests that manipulation of odours on a score such as verbalisability can still be useful for odour categorisation, as it is strongly related to the familiarity dimension that the analysis in

Chapter 2 has determined suitable for use. Indeed, when considered in the context of previous findings that have shown working memory performance differences based on a similar verbalisability score (e.g. Jönsson et al. 2011), and the findings in Chapters 3, 4, and 5 also showing memory differences based on these scores, there is support for use of this dimension in experimental control.

These normative data were used for stimulus control across a number of experiments in this thesis. A robust working memory advantage was observed for verbalisable odours controlled on intensity and hedonic strength dimensions. It should be noted, however, that controlling stimuli whilst maintaining the difference across the dimension of interest is not simple due to the correlations observed between familiarity, intensity, and pleasantness. That is, although there was a non-significant relationship between familiarity and intensity, there was a positive relationship between familiarity and pleasantness, and a negative relationship between intensity and pleasantness, which meant there was inevitably some trade-off when selecting odour sets; odours manipulated on familiarity but controlled on intensity would typically also show differences in pleasantness ratings. However, this experimental issue was mitigated somewhat by controlling differences in hedonic strength, which was a variable calculated from the odour's deviation in pleasantness from a neutral midpoint. This allowed low verbalisability sets to be more unpleasant than verbalisable sets, but the magnitude of the hedonic response to the odours was similar.

Together, despite some methodological challenges in its implementation in memory experiments, this thesis has demonstrated effective control and manipulation of the normative data associated to odorant stimuli.

6.2.2 Olfactory working memory ability

Chapters 3 and 4 have built on only two previous published studies that have examined olfactory working memory capabilities (Dade et al., 2001; Jönsson et al., 2011). Specifically, working memory is required for the manipulation (updating) of stored short-term representations of items and their position in a remembered sequence. Odour verbalisability improved n-back performance, replicating the finding in Jönsson et al. (2011) and further supporting suggestions that verbal information will facilitate odour memory (Jehl et al., 1997; Kärnekull et al., 2015; Lyman & McDaniel, 1986, 1990). Furthermore, like Jönsson et al., n-back performance for low verbalisability odours was above chance, supporting an updating process in odour memory that is not reliant on verbal information.

The key contribution of these chapters, however, were that this verbalisability advantage was unaffected by concurrent articulation (CA). This has provided evidence against verbal recoding and the subsequent use of a verbal rehearsal strategy in a working memory task. However, there was some effect of concurrent rotation (manifesting in an increase in false alarms), which provided evidence that an attention-demanding maintenance process was being performed in the retention interval. Furthermore, the advantage for high verbalisable odours appeared to be independent of perceptual familiarity (though see Section 6.2.2.3 for a discussion of this finding), and instead was due to a reduction in the ability to apply a controlled working memory process when odours were difficult to verbally label. Finally, Chapter 4 provided evidence that working memory performance for low verbalisability odours was unrelated to performance in verbal, visual, and verbalisable odour n-back tasks.

6.2.2.1 Verbal representation in olfactory working memory

A prosaic assumption for the n-back advantage for verbalisable odours is verbal recoding and subsequent application of a verbal rehearsal process for maintenance and updating. Indeed, models of n-back performance typically describe a rehearsal process for maintaining to-be-remembered items, which must be linked to their presentation order and updated as new items are presented (Chatham et al., 2011; Juvina & Taatgen, 2007; Ralph, 2014; Szmalec et al., 2011).

Whether this verbalisable advantage was due to the adoption of a verbal rehearsal strategy was explored in Experiment 3.2 using a dual-task procedure. Potential outcomes were that (1) CA would impair both low and high verbalisability odour performance, reflecting n-back performance that is based solely on rehearsal of labels that differ in quality, (2) that the verbalisable odours will be impaired by CA but leave the low verbalisability odours unaffected, reflecting a verbal rehearsal process for the verbalisable odours and some other non-verbal strategy for the low verbalisable odours, or (3) olfactory n-back performance will be unaffected by CA, reflecting an advantage for the verbalisable odours that is not due to verbal rehearsal. These data supported outcome (3), suggesting a verbal rehearsal strategy was not applied for maintenance and updating of odours during the n-back task. Whilst these findings should be considered with the caveat that a simple counting concurrent articulation task may not have been sufficient to impair a verbal rehearsal process (cf. Vuontela et al., 1999), it suggests that verbal rehearsal was not responsible for the observed verbalisable odour working memory advantage. Furthermore, the findings are proposed as evidence against a general reliance on verbal rehearsal for all odours in the task.

6.2.2.2 Non-verbal working memory processes

The use of controlled working memory processes (e.g. a maintenance and updating mechanism) in the n-back task, and the influence of familiarity-based responding (decisions based on an item strength signal through recent exposure), was investigated in Experiments 3.2, 3.3 and 4.1. The dual-task procedure in Experiment 3.2 additionally applied a concurrent mental rotation task, designed to load attentional resources that may otherwise be engaged in some maintenance and updating mechanism. The data provides some support for the proposition that there was a working memory strategy engaged for both verbalisable and low verbalisable odours that required attentional control in the stimulus retention interval. The evidence for this was not particularly convincing, however, as the effect of concurrent rotation affected false alarms rates but was not replicated for hits or *A'*. Furthermore, the effect on false alarms occurred as a general effect on all odours rather than an interaction across odour sets, suggesting little difference in the application of attention across low and high verbalisability odours. However, the findings were presented as tentative evidence for an attention refreshing mechanism that can be used for maintenance and manipulation in the n-back task. This attentional refreshing process is described as directed attention to activated items in memory (e.g. the to-be-remembered window of length *n* during the n-back task), and is proposed as an alternative maintenance process for both verbal and non-verbal stimuli in working memory (Baddeley, 2012; Camos et al., 2011).

The remember-know paradigm (Experiment 3.2) was amended to further assess the use of control strategies in the n-back task, and an individual-differences design (Experiment 4.1) was used to examine whether verbal and visual n-back performance (thought to be driven by controlled working memory resources) was related to low or high verbalisable odour n-back performance. The high proportion of recollection for verbalisable odours in Experiment 3.3 supported engagement of a control process for these odours, and a decrease in only *remember* responses for low verbalisability stimuli was interpreted as a decrease in successful application of this control process.

Recollection has been shown to increase with greater conceptual salience (Rajaram, 1998), and has been used to explain episodic memory differences for odours (M. J. Olsson et al., 2009). The increase in controlled working memory processing for familiar odours is also proposed to be the result of this increased saliency, which is associated to an increased ability to bind the remembered item to its context (Oberauer, 2005). That is, the ability to represent an odour representation is improved by the availability of semantic information, and this improvement acts specifically on the ability to engage updating processes on this representation. Indeed, conceptual information is proposed to be necessary in the formation of a mental image (Kan, Barsalou, Olseth Solomon, Minor, & Thompson-Schill, 2003), and in olfaction may support the retrieval of a stored representation in long-term memory in order to create a conscious perceptual image (Kosslyn, 2003).

This explanation was supported in Experiment 4.1, where n-back performance for low verbalisability odours was unrelated to verbal and visual n-back performance, whilst verbalisable odour n-back performance was related to both. Furthermore, there was only a weak positive correlation between the two odour tasks, and this was suggested to reflect individual differences in odour discriminability. Finally, there were no observed differences in discriminability across the familiar and unfamiliar odours sets, contradicting the proposition that the difference between high and low verbalisability odours simply reflected differences in perceptual discriminability (see below). Together with the evidence for a reduction in controlled processing in the remember-know task, the findings further suggest that availability of semantic information may be necessary for effective application of a controlled working memory strategy for odour stimuli.

6.2.2.3 Perceptual familiarity effects on olfactory n-back performance

A strong relationship between verbalisability and familiarity was observed in Chapter 2, such that odours classified as easy to verbalise also show strong ratings of familiarity. It was therefore investigated, given the null effect of CA in Experiment 3.2, whether the working memory advantage for high verbalisability odours was a result of enhanced discriminability through perceptual learning. Perceptual familiarity effects (i.e. an advantage from 'mere exposure') has been observed in several long-term olfactory memory experiments, and is proposed to underlie the mnemonic basis of object-processing accounts of olfactory perception (Stevenson & Wilson, 2007). This perceptual effect was assessed in Experiment 3.4 by attempting to induce familiarity in odours classified as low verbalisability.

Perceptual familiarity of odours with low normative familiarity was experimentally increased by having participants perform a number of rating and discrimination tasks (see Sinding et al., 2015 for a similar application of this design). Whilst it was predicted that increased perceptual familiarity would increase working memory performance, the opposite effect was found. Indeed, it was not only the familiarised odours that saw poorer performance, but performance in general (including for odours that had high normative familiarity and had not been experienced in the previous task) that was lower in the group pre-exposed to odours.

This study therefore failed to demonstrate a perceptual learning effect on olfactory working memory performance, and contrasts several findings that suggest such exposure is essential for odour discriminability (e.g. Rabin, 1988; Stevenson, 2001;

Case & Stevenson 2004). However, rather than contradicting these findings, the lack of perceptual learning could be explained by the following two reasons. First, the number of prior exposures may not have been sufficient. Indeed, Reder et al. (2015) presented stimuli across multiple sessions over several weeks, compared to the pre-exposure task here which included only 14 presentations in a single session. However, it should be noted that Sinding et al. (2015) presented odours to participants only 22 times, and the familiarisation sessions in Jehl et al. (1995) only presented each odour once in each of three sessions. Second, an increase in familiarity from the pre-exposure session may have affected the reliability of a familiarity signal. That is, the unreliability of such a signal meant participants were less likely to accept target items. In summary, the study supports an important contribution of a familiarity signal to n-back strategy and target judgements, but not the expected demonstration of perceptual learning through prior exposure (see also Jehl et al. 1997, for demonstration of response confusions in olfactory memory due to prior exposure).

6.2.2.4 *Familiarity-based processes in olfactory working memory*

If low verbalisability odours are characterised by an absence of controlled working memory resources, how is above-chance performance observed for these odours? It was suggested that working memory performance for low verbalisability odours was due to the contribution of some familiarity-based perceptual memory to an n-back decision. This process for recognising odours has been proposed in Wilson and Stevenson (2006), and has also been described as RWI (Cleary, 2002; Cleary et al., 2010). That is, recognition may be the consequence of residual activation from a perceptual representation, which is not consciously accessible, but can drive performance in memory tasks.

In the n-back task, perceptual-based recognition that employs a familiarity signal (i.e. residual activation) has been proposed to enable above-chance performance without requiring active rehearsal of stimuli (Juvina & Taatgen, 2007). Instead, the activation of memory traces is compared to an estimated 'time-tag' for a target, essentially allowing an n-back decision based on an educated assessment of familiarity. That is, the activation of the presented item is matched to the activation level that would be expected if this item was a target.

The findings in the present study are, however, equivocal over the use of such a low-control strategy for low verbalisable/unfamiliar odours. For example, Experiment 3.2 suggests that a familiarity-based strategy is not used because the inclusion of a concurrent rotation task saw a general detrimental effect on task performance (although this was only shown with false alarms). This effect would not be expected if n-back decisions were made based on an assessment of familiarity at probe presentation, so was consequently proposed as tentative evidence for a rehearsal mechanism in the n-back retention interval (i.e. attentional refreshing) for both low and high verbalisability odours. Furthermore, though there was a reduction in controlled working memory processes for low verbalisability odours (as measured by a reduction in *remember* responses), there was no corresponding increase in strategies based on experimental familiarity (as measured by *know* responses). This suggests that there is not application of a low-control strategy (i.e. a strategy that compares experimental familiarity of unfamiliar odours to a plausible familiarity level for a target), but instead suggests the successful application of a control strategy is reduced for most unfamiliar odours (as indexed by an increase in 'guess' responses).

However, there was no relationship in Experiment 4.1 between low verbalisability odours and the verbal and visual n-back tasks, indicating that the above-chance performance is driven by something other than simply reduced ability to recollect the position of the odours. That is, a relationship would still be expected if the n-back strategy was relying on working memory resources, albeit in a more difficult task. Indeed, n-back performance in the verbal 3-back task was similar to the low verbalisable odour 2-back tasks, but was still related to both visual and high verbalisability 2-back tasks. This suggests that performance for low verbalisability odours is unrelated to general working memory ability (utilised for verbal stimuli, visual stimuli, and verbalisable odours), and perhaps indicates a qualitatively different n-back strategy for low verbalisable/unfamiliar odours.

The proposed interpretation of the remember-know paradigm was a reduction in recollection without a corresponding reliance on familiarity-based processes for the low verbalisability odours. However, the findings in Chapter 4 contradict such an interpretation, as the lack of a relationship with established working memory measures suggest a qualitatively different strategy. A possible alternative explanation is discussed here, based on the dramatic increase in *guess* responses for low verbalisability odours. It is possible that participants did not show an increase in *know* responses for these low verbalisability odours because they struggled to distinguish such decisions from *guess* responses. Indeed, guess responses may have simply reflected a weak criterion of acceptance based on a familiarity-based strategy (see Gardiner et al., 1998 for a discussion of the use of guess reponses in the remember-know paradigm). Additional analysis of the remember-know findings by collapsing *guess* and *know* responses demonstrates a clear interaction, where responses shift to heavy familiarity-based (that is, a know/guess response) when odours were difficult to verbalise. Of course, such an interpretation should be treated with caution as a *guess* response was used as a catch-all for responses judged by the participants to have been made based on something other than memory. This proposal warrants further investigation however, perhaps by assessing the ability of guess responses to discriminate targets and lures. That is, if guess responses are able to generate more hits than false alarms, it would suggest that guess responses are made based on some residual memory trace.

Based on similar suggestions for a perceptually-based recognition process for unidentified odours (Cleary et al. 2010) that does not enable conscious imagery or rehearsal (Stevenson, 2009), it seems likely that a similar process has been observed for the low verbalisability odours in the present thesis. Recognition as a result of residual, perceptually-based, activation is able to drive above-chance performance in the n-back task. This is particularly the case if participants employed a *low-control* strategy of responding, where this activation was compared to a plausible level of familiarity for a target item.

6.2.3 Absent proactive interference in olfactory memory

Chapter 5 was the first study to date in which the recent probe PI procedure was applied to odours. In this study, low verbalisability odours did not show a recent-probes effect, typically used to demonstrate PI in other modalities. This is an interesting finding that contrasts previous evidence of strong proactive interference for odours (Lawless & Engen, 1977). In contrast, high verbalisability odours did show a PI effect, suggesting the verbalisable characteristic of odours produce proactive interference effects similar to other modalities. To rule out an absence of an effect due to methodological differences, in two further experiments the PI effect was demonstrated with verbal (word) stimuli, and faces. These findings provide a first assessment of item-specific PI effects in a short-term olfactory memory task, and further demonstrate that olfactory memory without accompanying semantic or verbal information may be qualitatively different to other modalities. This section attempts to accommodate these findings with the n-back findings presented above.

A familiarity-inhibition interpretation of the recent-probes effect was described in Chapter 5, where the strong familiarity that provokes a positive response conflicts with the memory of the TBR set. A weak link is proposed to exist between the low verbalisability odours and the familiarity signal present as a result of recent presentation. This weak link is thought to result from its low distinctiveness (or fuzzy representation, as described in Wilson & Stevenson 2006). Consequently, the conflicting 'yes' signal associated to that item is weaker, and results in the absent PI effects.

How then, might these findings converge with those demonstrated in the n-back task? The absent proactive interference would predict little effect of recent-lure familiarity on participant ability to correctly reject these items, if there is indeed a weak link between this conflicting 'yes' response and the probe item. This was not observed in Chapter 4, where greater close-lure false alarms were shown for the low verbalisability odours. Consequently, a weak item-to-familiarity link may be unrelated to performance differences in the n-back task, which are proposed due to application of differing working memory resources. Indeed, it should be noted that using a familiarity signal to make an n-back decision is a noisy method of making a judgement (Juvina & Taatgen, 2007). It is therefore possible that the controlled retrieval process applied for verbalisable odours in the n-back task is susceptible to proactive interference effects, but the detriment to performance as a result of reliance on a familiarity-based process for low verbalisability odours masks this effect. Further research, perhaps with a more controlled manipulation of recent-lure items in the n-back tasks (see below), may elucidate the possible independent effects of PI and working memory processes on n-back performance.

6.3 Accommodation within olfactory-specific or general models of working memory

6.3.1 Olfactory-centred unitary model

Several findings in this thesis are consistent with the specifications of a pattern-matching process for odour perception and memory (Stevenson & Boakes, 2003; Stevenson & Mahmut, 2013; Stevenson & Wilson, 2007; Wilson & Stevenson, 2006). First, low verbalisability odours are proposed to be difficult to verbalise because the outcome of pattern matching results in weak activation of multiple stored engrams (Stevenson & Boakes, 2003). This ‘fuzzy’ percept for low verbalisability odours (Stevenson & Mahmut, 2013, p. 1428) may be responsible for the absent PI effects observed in Chapter 5.

Second, the model proposes short-term memory for odours to be the result of residual activation of an odour engram following initial presentation of the odour, and that this activation is not available to consciousness. That is, the perceptual representation is suggested to be remembered by assessing its activation upon re-presentation of the same odour, in a process similar to recognition without identification (Cleary et al., 2010). Such a representation would not be available to maintenance or updating in an n-back working memory task, and for low verbalisability odours the evidence seems to support this.

Second, Stevenson and colleagues suggest that naming and other available semantic information serve memory by providing an additional means with which the odour may be remembered. This thesis finds no evidence that a verbal rehearsal mechanism is used to maintain a verbal label for an odour. However, the evidence presented suggests that the availability of semantic information allows engagement of working memory resources (cf. the general effect of concurrent rotation in Experiment 3.2). A refreshing mechanism is proposed for the maintenance of odours, which requires controlled attention applied to a stored representation. Furthermore, conceptual knowledge is suggested to be a necessary component for conscious access to this internal representation, or olfactory image, to occur (Kan et al., 2003; Tomiczek & Stevenson, 2009). Consequently, contrary to the suggestion that olfactory imagery is not available to consciousness, an internal perceptual odour representation may be used for the purposes of maintenance, but appears to require conceptual knowledge to allow reactivation of the stored long-term representation (Kosslyn, 2003; Rinck et al., 2009).

6.3.2 Modularity

The multicomponent memory framework (e.g. Baddeley & Hitch, 1974; Baddeley, 2000) describes specialised processes for maintenance and manipulation of information in working memory. The findings in this thesis have demonstrated unique characteristics of low verbalisability odour memory that are thought to reflect memory for a perceptually-based odour representation that can occur within an olfactory buffer (e.g. Andrade & Donaldson 2007; Zelano et al., 2009). Indeed, there are characteristics of this olfactory memory that suggest it may qualitatively differ to verbal and visuo-spatial subsystems.

Previous inconsistent findings in respect to whether olfactory memory is qualitatively different to other stimulus types may be based on stimulus selection. That is, the characteristics of odours selected for research, particularly their identifiability, has been proposed to affect the way odours are represented in memory (Zelano et al., 2009). Zelano et al. suggest that unidentified odour memory reflects perceptual processing of odours in an independent olfactory buffer.

Chapter 5 has shown an absence of PI for low verbalisability odours, presumed to reflect olfactory memory based on a perceptual representation. Although the discussion above has focussed on amodal explanations that consider the saliency of odours and a possible poor link between an odour and its familiarity signal, such findings may also be accommodated within a modular working memory system that describes a subsystem qualitatively different to other memory systems.

An important finding in this thesis, and one which has been demonstrated in some short-term memory tasks (Andrade & Donaldson, 2007), is that olfactory working memory does not require a verbal rehearsal process for above chance performance to occur. This suggests a perceptual representation that can be applied in a working memory task that is analogous to other modalities, though it appears that there is an important role of semantic information to allow utilisation of working memory resources similar to those seen for other modality tasks. A perceptual representation with little conceptual salience may instead see reliance on a familiarity-based process that requires minimal, if any, working memory resources. Such a finding suggests any representation in an olfactory buffer interacts with long-term semantic information before executive functions can act upon the item in memory.

An olfactory slave system could conceivably take on many of the properties described in Wilson and Stevenson (2006), such as a pattern-matching process, and communication with verbal information and long-term memory to establish an odour object. Indeed, as noted by Baddeley (2012) and Logie (2011), the processes in the multicomponent model of working memory are not too far removed from the unitary model described in Cowan (1999). That is, the focus of attention that activates representations in LTM can be considered similar to executive resources acting upon a representation within the episodic buffer. Such an interaction between slave systems and stored semantic information would not be unique to olfaction, as it is also proposed in the multicomponent model for both visual and verbal working memory (see Baddeley, 2012). Furthermore, the episodic buffer has been proposed as the system that enables access to consciousness. Considering the findings in the present thesis that proposed a link between semantic information and the ability to create an internal representation of an odour, working memory performance for verbalisable odours may reflect executive functions acting up representations in the episodic buffer.

6.4 Limitations and further research

A potential criticism of the application of the n-back procedures is the poor control of recent-lure items. Although the n-back tasks in Chapter 3 included close-lure trials, these were not consistent between participants, and were typically fewer than the number of targets. As described in Ralph (2014), the presence of these recent- and non-recent lures may affect the strategy employed during the n-back task. That is, a low number of recent-lures might, after several trials where they establish the type of items being presented, might result in a participant deciding that the reward for a high-control strategy is too low, when a familiarity-based strategy can be used to get a high percentage of trials correct. However, the low number of trials in these olfactory tasks compared to verbal versions of the n-back task is likely to mean participants did not have enough time to effectively gauge the proportion of recent-lures to targets. Furthermore, at least for the high verbalisability odours, the remember-know findings suggest that recollection-based processes are being applied.

The high number of non-recent lure items may have had another effect. Whilst there was no bias in the number of recent and non-recent lures across low and high odour verbalisability conditions, issues may arise from the use of non-recent lures when establishing the level of n-back performance. The inclusion of these lures (which can be rejected on the basis of familiarity alone) in an index of n-back performance inflates this estimate of performance. This is because these rejections do not require the controlled processes prescribed in working memory. This is not problematic for the comparison of performance between low and high verbalisability odours because both tasks included the same number of these lures. However, when determining whether low-verbalisability n-back performance was above chance, it is difficult to differentiate performance based on automatic processes. Above chance performance for the low verbalisable odours may therefore be due to performance on the non-recent lures (an issue that is also pertinent for Jönsson et al., 2011). Indeed, as demonstrated, the effect of these automatic processes is particularly important in n-back performance for low verbalisability odours.

These issues were addressed to some extent in Experiment 4.1, where close-lures were controlled so that they were equal across participants and a similar number to the number of targets. Furthermore, these recent-lures were also used in the calculation of *A'* sensitivity, allowing an accurate assessment of controlled working memory ability that was not inflated by non-recent lure rejections requiring no recollection. Importantly, performance remained above chance for low verbalisable odours despite the removal of non-recent lures from the analysis. However, it should be noted that if participants are employing a familiarity-based strategy using a single criterion for judging an n-back target, slightly above-chance performance would still be expected due to the use of recent-lures that are outside the n-back window. That is, a judgement of familiarity after setting a criterion at the *n*-back position is likely sensitive enough, at least some of the time, to discriminate between a target item and a recent-lure from $n + 1$ (i.e. the item is slightly less temporally familiar than the target). The findings presented above would therefore be well-served by a systematic assessment of recent-lure effects on working memory performance, and in particular should consider the use of a working memory performance index based on targets and only *n minus* recent lures, as this would mean all close lures would fall above the response criterion. It may be necessary, then, to explore olfactory n-back performance with a more difficult 3-back task to allow a more varied manipulation of lure types in the task (although whether this would produce above chance performance is an empirically testable question).

An additional debate concerns the stimulus choice in the present experiment. The olfactory stimuli used in this thesis have utility in olfactory memory experiments due to the normative data established in Chapter 2, and because they are commercially and easily available for researchers. However, greater control of the odours may be required in studies interested in the chemical complexity of odours, and studies that wish to control the intensity of a single odour would have difficulty doing so without knowledge of the solutions that make up these stimuli; information which is not available. Potential future research into this corpus of odour stimuli would benefit from a detailed assessment of odour similarities, perhaps presented as a matrix of similarity to all other odours. Such analysis of 200 odours would be a huge undertaking, however. In addition, a cross-cultural application of the method in Chapter 2 would serve to diversify the normative data, and allow odour sets to be tailored to participant individual differences and their likely prior experience without requiring assessment of the odours by the participant themselves.

Despite the above discussion points regarding the olfactory stimuli used, future research into olfactory memory would benefit from the rigour/control applied to odours in the present experiment. For example, serial-recall has been used to support qualitative differences in olfactory memory (A. J. Johnson & Miles, 2009), though alternative explanations such as the SIMPLE model suggest differences may be dependent on the psychological distinctiveness of these stimuli (Hay et al., 2007). Such predictions are empirically testable by controlling dimensions using the present corpus of odours, and may serve to elucidate the serial-position functions observed between olfactory and other-modality stimuli.

6.5 Summary and conclusion

This thesis has examined the ability to represent odours in working memory, and explored qualitative differences in memory when odours differ in available semantic information (measured by verbalisability). It contributes to the study of olfactory memory by presenting a normative database for odour stimuli, and showing that whilst individual differences are an important consideration when selecting odours based upon a perceptual dimension, differences between participants are small enough to discriminate differences based on the features of an odour (Chapter 2). A working memory performance advantage was observed for verbalisable odours that replicates previous findings (Jönsson et al. 2011), but this was not due to perceptual learning or verbal rehearsal (Chapter 3). Indeed, odours low on verbalisability showed reduced application of controlled working memory processes (Chapter 3), and working memory performance was unrelated to other working memory tasks including verbal and visual stimuli (Chapter 4). Finally, low verbalisability odours showed no susceptibility to proactive interference in a recent-probes task, compared to high verbalisability odours, visual stimuli, and verbal stimuli (Chapter 5).

The working memory findings may be accommodated in an amended olfactory-centred unitary model (Wilson & Stevenson 2006), where a perceptually-represented odour is recognised by the residual activation of a stored olfactory pattern. The access of consciousness to this representation is unavailable (see Stevenson, 2009), meaning working memory processes such as maintenance and updating cannot be performed.

However, when semantic information is available, a refreshing process (Baddeley, 2012; M. R. Johnson et al., 2015; Raye et al., 2007) is proposed to be available for the maintenance and updating requirements in an n-back task, achieved because the conceptual salience of these items allow conscious access to an internal representation to be achieved (Kan et al., 2003).

7 References

Allen, R. J., Hitch, G. J., Mate, J., & Baddeley, A. D. (2012). Feature binding and attention in working memory: A resolution of previous contradictory findings. *The Quarterly Journal of Experimental Psychology*, *65*(March 2015), 2369–2383. https://doi.org/10.1080/17470218.2012.687384

Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc*, *31*(4), 578–602. https://doi.org/10.3758/BF03200738

Anderson, J. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261–295. Retrieved from http://www.sciencedirect.com/science/article/pii/S0022537183902013

Andrade, J., & Donaldson, L. (2007). Evidence for an Olfactory Store in Working Memory? *Psychologia*, *50*(2), 76–89.

Annett, J. M., Cook, N. M., & Leslie, J. C. (1995). Interference with olfactory memory by visual and verbal tasks. *Perceptual and Motor Skills*, *80*(3 Pt 2), 1307–1317. https://doi.org/10.2466/pms.1995.80.3c.1307

Annett, J. M., & Leslie, J. C. (1996). Effects of visual and verbal interference tasks on olfactory memory: the role of task complexity. *British Journal of Psychology*, *87*(3), 447–460.

Annett, J. M., & Lorimer, A. W. (1995). Primacy and recency in recognition of odours and recall of odour names. *Perceptual and Motor Skills*, *81*(3 (pt. 1)), 787–794. https://doi.org/10.2466/pms.1995.81.3.787

Arshamian, A., & Larsson, M. (2014). Same same but different: The case of olfactory imagery. *Frontiers in Psychology*, *5*(FEB), 34. https://doi.org/10.3389/fpsyg.2014.00034

Arshamian, A., Olofsson, J. K., Jönsson, F. U., & Larsson, M. (2008). Sniff your way to clarity: The case of olfactory imagery. *Chemosensory Perception*, *1*(4), 242–246. https://doi.org/10.1007/s12078-008-9035-z

Ashton, R., & White, K. D. (1980). Sex differences in imagery vividness: An artifact of the test. *British Journal of Psychology*, *71*(1), 35–38. https://doi.org/10.1111/j.2044-8295.1980.tb02726.x

Atkins, A. S., Berman, M. G., Reuter-Lorenz, P. a, Lewis, R. L., & Jonides, J. (2011). Resolving semantic and proactive interference in memory over the short-term. *Memory & Cognition*, *39*(5), 806–817. https://doi.org/10.3758/s13421-011-0072-5

Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A Proposed System and its Control Processes. *Psychology of Learning and Motivation*, *2*, 89–195. https://doi.org/10.1016/S0079-7421(08)60422-3

Attneave, F., Arnoult, M. D., & Attneave, F. (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin*, *53*(6), 452–471. https://doi.org/10.1037/h0044049

Avons, S. E. (1998). Serial report and item recognition of novel visual patterns. *British Journal of Psychology*, *89*(2), 285–308. https://doi.org/10.1111/j.2044-8295.1998.tb02685.x

Awh, E., Jonides, J., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology. Human Perception and Performance*, *24*(3), 780–90. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9627416

Ayabe-Kanamura, S., Schicker, I., Laska, M., Hudson, R., Distel, H., Kobayakawa, T., & Saito, S. (1998). Differences in perception of everyday odors: a Japanese-German cross- cultural study. *Chemical Senses*, *23*(1), 31–38. https://doi.org/10.1093/chemse/23.1.31

Baddeley, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *The Quarterly Journal of Experimental Psychology*, *18*(912873514), 302–309. https://doi.org/10.1080/14640746608400047

Baddeley, A. D. (1986). *Working Memory*. Oxford, UK: Oxford University Press.

Baddeley, A. D. (1996). The fractionation of working memory. *Proceedings of the National Academy of Sciences*, *93*(24), 13468–13472. https://doi.org/10.1073/pnas.93.24.13468

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. D. (2007). *Oxford Psychology Series: Working memory, Thought, and Action*. Oxford, UK: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198528012.001.0001

Baddeley, A. D. (2012). Working Memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Baddeley, A. D., Allen, R. J., & Hitch, G. J. (2011). Binding in visual working memory: The role of the episodic buffer. *Neuropsychologia*, *49*(6), 1393–1400. https://doi.org/10.1016/j.neuropsychologia.2010.12.042

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology. General*, *129*(1), 126–45. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10756490

Baddeley, A. D., & Hitch, G. J. (1974). *Working memory*. (G. H. Bower, Ed.), *The psychology of learning and ...* (Vol. VIII). Oxford, UK: Clarendon Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A. D., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology*, *36A*(2), 233–252. https://doi.org/10.1080/14640748408402157

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 575–589. https://doi.org/10.1016/S0022-5371(75)80045-4

Badre, D., & Wagner, A. D. (2005). Frontal lobe mechanisms that resolve proactive interference. *Cerebral Cortex*, *15*(12), 2003–2012. https://doi.org/10.1093/cercor/bhi075

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual Word Recognition of Single- Syllable Words. *Journal of Experimental Psychology: General*, *133*(2), 283–316.

Banich, M. T. (2009). Executive function: The search for an integrated account. *Current Directions in Psychological Science*, *18*(2), 89–94. https://doi.org/10.1111/j.1467-8721.2009.01615.x

Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual Differences in Working Memory Capacity and Dual-Process Theories of the Mind. *Psychological Bulletin*, *130*(4), 553–573. https://doi.org/10.1037/0033-2909.130.4.553

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time Constraints and Resource Sharing in Adults' Working Memory Spans. *Journal of Experimental Psychology: General*, *133*(1), 83–100. https://doi.org/10.1037/0096-3445.133.1.83

Barrouillet, P., & Lecas, J.-F. (1999). Mental Models in Conditional Reasoning and Working Memory. *Thinking & Reasoning*, *5*(4), 289–302. https://doi.org/10.1080/135467899393940

Bensafi, M. (2002). Autonomic Nervous System Responses to Odours: the Role of Pleasantness and Arousal. *Chemical Senses*, *27*(8), 703–709. https://doi.org/10.1093/chemse/27.8.703

Bensafi, M., Sobel, N., & Khan, R. M. (2007). Hedonic-specific activity in piriform cortex during odor imagery mimics that during odor perception. *Journal of Neurophysiology*, *98*(6), 3254–3262. https://doi.org/10.1152/jn.00349.2007

Berglund, B., Berglund, U., Ekman, G., & Engen, T. (1971). Individual psychophysical functions for 28 odorants. *Perception & Psychophysics*, *9*(3), 379–384. https://doi.org/10.3758/BF03212672

Bhatarah, P., Ward, G., & Tan, L. (2006). Examining the relationship between free recall and immediate serial recall: The effect of concurrent task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 215–229. https://doi.org/10.1037/0278-7393.32.2.215

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, *6*(2), 173–189. https://doi.org/10.1016/0010-0285(74)90009-7

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol Review*, *108*(3), 624–652. https://doi.org/10.1037/0033-295X.108.3.624

Bradley, M., & Lang, P. J. (1994). Measuring Emotion: The Self-Assessment Semantic Differential Manikin and the. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(I), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Bradshaw, G. L., & Anderson, J. R. (1982). Elaborative encoding as an explanation of levels of processing. *Journal of Verbal Learning and Verbal Behavior*, *21*(2), 165–174. https://doi.org/10.1016/S0022-5371(82)90531-X

Brandon, M., Jha, A. P., Trueswell, J. C., Barde, L. H. F., & Thompson-schill, S. L. (2003). Proactive Interference in Verbal and Non-Verbal Working Memory. In *the Annual Meeting of the Psychonomic Society*. Vancouver, Canada.

Braver, T. S. (2012, February). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2011.12.010

Brewer, W. F., & Sampaio, C. (2012). The metamemory approach to confidence: A test using semantic memory. *Journal of Memory and Language*, *67*(1), 59–77. https://doi.org/10.1016/j.jml.2012.04.002

Bridson, N. C., Muthukumaraswamy, S. D., Singh, K. D., & Wilding, E. L. (2009). Magnetoencephalographic correlates of processes supporting long-term memory judgments. *Brain Research*, *1283*, 73–83. https://doi.org/10.1016/j.brainres.2009.05.093

Broadbent, D. (1958). *Perception and communication*. London: Pergamon Press. https://doi.org/10.1108/eb015727

Brown, G. D. a., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory & Cognition*, *15*(3), 208–216. https://doi.org/10.3758/BF03197718

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–11. https://doi.org/10.3758/s13428-013-0403-5

Buck, L. B. (2000). The molecular architecture of odor and pheromone sensing in mammals. *Cell*, *100*(6), 611–618. https://doi.org/10.1016/S0092-8674(00)80698-4

Buck, L. B. (2004). Olfactory receptors and odor coding in mammals. *Nutrition Reviews*, *62*(11 Pt 2), S184-NaN-S241. https://doi.org/10.1301/nr.2004.nov.S184

Cain, W. S. (1969). Odor intensity: Differences in the exponent of the psychophysical function. *Perception & Psychophysics*, *6*(6), 349–354. https://doi.org/10.3758/BF03212789

Cain, W. S. (1979). To know with the nose: keys to odor identification. *Science*, *203*(4379), 467–470. https://doi.org/10.1126/science.760202

Cain, W. S., De Wijk, R. A., Lulejian, C., Schiet, F., & See, L.-C. (1998). Odor Identification: Perceptual and Semantic Dimensions. *Chemical Senses*, *23*(3), 309–326. https://doi.org/10.1093/chemse/23.3.309

Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, *61*(3), 457–469. https://doi.org/10.1016/j.jml.2009.06.002

Camos, V., Mora, G., & Oberauer, K. (2011). Adaptive choice between articulatory rehearsal and attentional refreshing in verbal working memory. *Memory & Cognition*, *39*(2), 231–44. https://doi.org/10.3758/s13421-010-0011-x

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*(3), 404–431. https://doi.org/10.1037/0033-

295X.97.3.404

Carroll, B., Richardson, J. T. E., & Thompson, P. (1993). Olfactory Information Processing and Temporal Lobe Epilepsy. *Brain and Cognition*, *22*(2), 230–243. https://doi.org/10.1006/brcg.1993.1036

Cedrus. (2015). Superlab. Cedrus Corporation. Retrieved from http://cedrus.com

Cessna, T. C., & Frank, R. A. (2013). Does odor knowledge or an odor naming strategy mediate the relationship between odor naming and recognition memory? *Chemosensory Perception*, *6*(1), 36–44. https://doi.org/10.1007/s12078-013-9139-y

Chatham, C. H., Herd, S. A., Brant, A. M., Hazy, T. E., Miyake, A., O'Reilly, R., & Friedman, N. P. (2011). From an executive network to executive control: a computational model of the n-back task. *Journal of Cognitive Neuroscience*, *23*(11), 3598–619. https://doi.org/10.1162/jocn_a_00047

Chein, J. M., & Fiez, J. A. (2010). Evaluating models of working memory through the effects of concurrent irrelevant information. *Journal of Experimental Psychology. General*, *139*(1), 117–37. https://doi.org/10.1037/a0018200

Chu, S., & Downes, J. J. (2000). Odour-evoked Autobiographical Memories: Psychological Investigations of Proustian Phenomena. *Chemical Senses*, *25*(1), 111–116. https://doi.org/10.1093/chemse/25.1.111

Chuah, Y. M. L., Maybery, M. T., & Fox, A. M. (2004). The long-term effects of mild head injury on short-term memory for visual form, spatial location, and their conjunction in well-functioning university students. *Brain and Cognition*, *56*(3), 304–312. https://doi.org/10.1016/j.bandc.2004.08.002

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A Taxonomy of External and Internal Attention. *Annual Review of Psychology*, *62*(1), 73–101. https://doi.org/10.1146/annurev.psych.093008.100427

Cleary, A. M. (2002). Recognition with and without identification: dissociative effects of meaningful encoding. *Memory & Cognition*, *30*(5), 758–767. https://doi.org/10.3758/BF03196431

Cleary, A. M., Konkel, K. E., Nomi, J. S., & McCabe, D. P. (2010). Odor recognition without identification. *Memory & Cognition*, *38*(4), 452–460. https://doi.org/10.3758/MC.38.4.452

Cleary, A. M., Winfield, M. M., & Kostic, B. (2007). Auditory recognition without identification. *Memory & Cognition*, *35*(8), 1869–1877. https://doi.org/10.3758/BF03192921

Cocchini, G., Logie, R. H., Della Sala, S., MacPherson, S. E., & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: evidence for domain-specific working memory systems. *Memory & Cognition*, *30*(7), 1086–95. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12507373

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, *386*, 604–608. Retrieved from http://www.academia.edu/download/38073548/Cohen-Nature_fMRI_WM.pdf

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. Retrieved from http://psycnet.apa.org/journals/rev/82/6/407/

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A*, *33*(4), 497–505. https://doi.org/10.1080/14640748108400805

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, *55*(1), 75–84. https://doi.org/10.1111/j.2044-8295.1964.tb00899.x

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity , short-term memory capacity , processing speed , and general f luid intelligence. *Intelligence*, *30*, 163–183. Retrieved from http://www.sciencedirect.com/science/article/pii/S0160289601000964

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003, December). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2003.10.005

Cowan, N. (1992). Verbal memory span and the timing of spoken recall. *Journal of Memory and Language*, *31*(5), 668–684. https://doi.org/10.1016/0749-596X(92)90034-U

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Mainenance and Executive Control* (pp. 62–101). Cambridge University Press.

https://doi.org/10.1017/S0140525X01003922

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *The Behavioral and Brain Sciences*, *24*(1), 87-114–85. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11515286

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, *169*, 323–38. https://doi.org/10.1016/S0079-6123(07)00020-9

Cowan, N. (2010). Multiple concurrent thoughts: The meaning and developmental neuropsychology of working memory. *Developmental Neuropsychology*, *35*(5), 447–474. https://doi.org/10.1080/87565641.2010.494985

Cowan, N., Elliott, E. M., Saults, S. J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005, August). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*. https://doi.org/10.1016/j.cogpsych.2004.12.001

Craig, K. S., Berman, M. G., Jonides, J., & Lustig, C. (2013). Escaping the recent past: Which stimulus dimensions influence proactive interference? *Memory & Cognition*, *41*(5), 650–670. https://doi.org/10.3758/s13421-012-0287-0

Crowder, R. G., & Schab, F. R. (1995). Imagery for odors. In R. G. Crowder & F. R. Schab (Eds.), *Memory for Odors* (pp. 93–107). Hillsdale, NJ: Erlbaum.

Curtis, C. E., & D'Esposito, M. (2003, September). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*.

https://doi.org/10.1016/S1364-6613(03)00197-9

Dacremont, C., & Valentin, D. (2004). Mobilization of short-term memory capacity for odors in discriminative tests: Implication for assessors' selection. *Food Quality and Preference*, *15*(7–8 SPEC.ISS.), 735–742. https://doi.org/10.1016/j.foodqual.2004.02.005

Dade, L. A., Zatorre, R. J., Evans, A. C., & Jones-Gotman, M. (2001). Working memory in another dimension: functional imaging of human olfactory working memory. *NeuroImage*, *14*(3), 650–60. https://doi.org/10.1006/nimg.2001.0868

Dalton, P., Maute, C., Oshida, A., Hikichi, S., & Izumi, Y. (2008). The use of semantic differential scaling to define the multidimensional representation of odors. *Journal of Sensory Studies*, *23*(4), 485–497. https://doi.org/10.1111/j.1745-459X.2008.00167.x

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory during reading. *Journal Of Verbal Learning And Verbal Behavior*, *19*(4), 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Davis, C. J. (2005). N-watch: a program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*(1), 65–70. https://doi.org/10.3758/BF03206399

De Araujo, I. E., Rolls, E. T., Velazco, M. I., Margot, C., & Cayeux, I. (2005). Cognitive modulation of olfactory processing. *Neuron*, *46*(4), 671–679. https://doi.org/10.1016/j.neuron.2005.04.021

De Bourbon-Teles, J., Bentley, P., Koshino, S., Shah, K., Dutta, A., Malhotra, P., … Soto, D. (2014). Thalamic control of human attention driven by memory and learning. *Current Biology*, *24*(9), 993–999. https://doi.org/10.1016/j.cub.2014.03.024

De Wijk, R. A., & Cain, W. S. (1994a). Odor identification by name and by edibility: Life-span development and safety. *Human Factors*, *36*(1), 182–187. https://doi.org/n.a.

De Wijk, R. A., & Cain, W. S. (1994b). Odor quality: Discrimination versus free and cued identification. *Perception & Psychophysics*, *56*(1), 12–18. https://doi.org/10.3758/BF03211686

Degel, J., Piper, D., & Köster, E. P. (2001). Implicit Learning and Implicit Memory for Odors: the Influence of Odor Identification and Retention Time. *Chemical Senses*, *26*(3), 267–280. https://doi.org/10.1093/chemse/26.3.267

Delon-Martin, C., Plailly, J., Fonlupt, P., Veyrac, A., & Royet, J. P. (2013). Perfumers' expertise induces structural reorganization in olfactory brain regions. *NeuroImage*, *68*, 55–62. https://doi.org/10.1016/j.neuroimage.2012.11.044

Depoorter, A., & Vandierendonck, A. (2009). Evidence for modality-independent order coding in working memory. *Quarterly Journal of Experimental Psychology (2006)*, *62*(3), 531–49. https://doi.org/10.1080/17470210801995002

Diana, R. A., & Reder, L. M. (2006). The low-frequency encoding disadvantage: Word frequency affects processing demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 805–815. https://doi.org/10.1037/0278-

7393.32.4.805

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*(July), 1–17. https://doi.org/10.3389/fpsyg.2014.00781

Distel, H., Ayabe-Kanamura, S., Martínez-Gómez, M., Schicker, I., Kobayakawa, T., Saito, S., & Hudson, R. (1999). Perception of Everyday Odors: Correlation between Intensity, Familiarity and Strength of Hedonic Judgement. *Chemical Senses*, *24*(2), 191–199. https://doi.org/10.1093/chemse/24.2.191

Distel, H., & Hudson, R. (2001). Judgement of odor intensity is influenced by subjects' knowledge of the odor source. *Chemical Senses*, *26*(3), 247–51. https://doi.org/10.1093/CHEMSE/26.3.247

Djordjevic, J., Zatorre, R. J., Petrides, M., Boyle, J. A., & Jones-Gotman, M. (2005). Functional neuroimaging of odor imagery. *NeuroImage*, *24*(3), 791–801. https://doi.org/10.1016/j.neuroimage.2004.09.035

Dong, S., Reder, L. M., Yao, Y., Liu, Y., & Chen, F. (2015). Individual differences in working memory capacity are reflected in different ERP and EEG patterns to task difficulty. *Brain Research*, *1616*, 146–156. https://doi.org/10.1016/j.brainres.2015.05.003

Doty, R. L., & Cameron, E. L. (2009). Sex differences and reproductive hormone influences on human odor perception. *Physiology and Behavior*, *97*(2), 213–228. https://doi.org/10.1016/j.physbeh.2009.02.032

Doty, R. L., Shaman, P., Applebaum, S. L., Giberson, R., Siksorski, L., & Rosenberg, L.

(1984). Smell identification ability: changes with age. *Science (New York, N.Y.)*, *226*(4681), 1441–1443. https://doi.org/10.1126/science.6505700

Doty, R. L., Shaman, P., Kimmelman, C. P., & Dann, M. S. (1984). University of Pennsylvania Smell Identification Test: a rapid quantitative olfactory function test for the clinic. *The Laryngoscope*, *94*(2 Pt 1), 176–8. https://doi.org/10.1288/00005537-198402000-00004

Doty, R. L., Snyder, P. J., Huggins, G. R., & Lowry, L. D. (1981). Endocrine, cardiovascular, and psychological correlates of olfactory sensitivity changes during the human menstrual cycle. *Journal of Comparative and Physiological Psychology*, *95*(1), 45–60. https://doi.org/10.1037/h0077755

Downing, P. E. (2000). Interactions between visual working memory and selective attention. *Psychological Science*, *11*(6), 467–473. https://doi.org/10.1111/1467-9280.00290

Dudukovic, N. M., Dubrow, S., & Wagner, A. D. (2009). Attention during memory retrieval enhances future remembering. *Memory & Cognition*, *37*(7), 953–61. https://doi.org/10.3758/MC.37.7.953

Duff, S. C., & Logie, R. H. (2001). Processing and storage in working memory span. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *54*(1), 31–48. https://doi.org/10.1080/02724980042000011

Dunn, J. C. (2004). Remember-Know: A Matter of Confidence. *Psychological Review*, *111*(2), 524–542. https://doi.org/10.1037/0033-295X.111.2.524

Dunn, J. C. (2008). The dimensionality of the remember-know task: a state-trace analysis. *Psychological Review*, *115*(2), 426–446. https://doi.org/10.1037/0033-295X.115.2.426

Ebbinghaus, H. (1885). *Űber das Gedächitnis: Unterschungen zur experimentellen Psychologie.* Duncker & Humblot.

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES--A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, *42*(1), 351–62. https://doi.org/10.3758/BRM.42.1.351

Eichenbaum, H., Morton, T. H., Potter, H., & Corkin, S. (1983). Selective olfactory deficits in case H.M. *Brain*, *106*(2), 459–472. https://doi.org/10.1093/brain/106.2.459

Elsley, J. V., & Parmentier, F. B. R. (2009). Is verbal-spatial binding in working memory impaired by a concurrent memory load? *Quarterly Journal of Experimental Psychology*, *62*(9), 1696–705.

Engen, T. (1982). *The perception of odors*. Academic Press.

Engen, T., Kuisma, J. E., & Eimas, P. D. (1973). Short-term memory of odors. *Journal of Experimental Psychology*, *99*(2), 222–225. https://doi.org/10.1037/h0034645

Engen, T., & Ross, B. M. (1973). Long-term memory of odors with and without verbal descriptions. *Journal of Experimental Psychology*, *100*(2), 221–227. https://doi.org/10.1037/h0035492

Engle, R. W., & Kane, M. (2004). Executive Attention, Working Memory Capacity, and a Two-Factor Theory of Cognitive Control. *The Psychology of Learning and Motivation: Advances in Research and Theory*, 145–199. https://doi.org/10.1016/S0079-7421(03)44005-X

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309–331. https://doi.org/10.1037/0096-3445.128.3.309

Evans, L. H., & Wilding, E. L. (2012). Recollection and Familiarity Make Independent Contributions to Memory Judgments. *Journal of Neuroscience*, *32*(21), 7253–7257. https://doi.org/10.1523/JNEUROSCI.6396-11.2012

Farah, M. J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review*, *95*(3), 307–317. https://doi.org/10.1037/0033-295X.95.3.307

Forster, K. I., & Bednall, E. S. (1976). Terminating and exhaustive search in lexical access. *Memory & Cognition*, *4*(1), 53–61. https://doi.org/10.3758/BF03213255

Fortin, N. J., Wright, S. P., & Eichenbaum, H. (2004). Recollection-like memory retrieval in rats is dependent on the hippocampus. *Nature*, *431*(7005), 188–191. https://doi.org/10.1038/nature02853

Frank, M. E., Goyert, H. F., & Hettinger, T. P. (2010). Time and intensity factors in identification of components of odor mixtures. *Chemical Senses*, *35*(9), 777–787. https://doi.org/10.1093/chemse/bjq078

Frank, R. A., Rybalsky, K., Brearton, M., & Mannea, E. (2011). Odor recognition memory as a function of odor-naming performance. *Chemical Senses*, *36*(1), 29–41. https://doi.org/10.1093/chemse/bjq095

Frasnelli, J., Charbonneau, G., Collignon, O., & Lepore, F. (2009). Odor localization and sniffing. *Chemical Senses*, *34*(2), 139–144. https://doi.org/10.1093/chemse/bjn068

Freiherr, J., Gordon, A. R., Alden, E. C., Ponting, A. L., Hernandez, M. F., Boesveldt, S., & Lundström, J. N. (2012). The 40-item Monell Extended Sniffin' Sticks Identification Test (MONEX-40). *Journal of Neuroscience Methods*, *205*(1), 10–16. https://doi.org/10.1016/j.jneumeth.2011.12.004

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*(2), 201–225. https://doi.org/10.1037/0096-3445.137.2.201

Ganis, G., & Kievit, R. (2015). *A New Set of Three-Dimensional Shapes for Investigating Mental Rotation Processes: Validation Data and Stimulus Set*. *Journal of Open Psychology Data* (Vol. 3). https://doi.org/10.5334/jopd.ai

Gardiner, J. M., Java, R. I., & Richardson-Klavehn, A. (1996). How level of processing really influences awareness in recognition memory. *Canadian Journal of*

Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, *50*(1), 114–122. https://doi.org/10.1037/1196-1961.50.1.114

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (1998). Experiences of remembering, knowing, and guessing. *Consciousness and Cognition*, *7*(7), 1–26. https://doi.org/10.1006/ccog.1997.0321

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (2002). Recognition memory and decision processes: a meta-analysis of remember, know, and guess responses. *Memory (Hove, England)*, *10*(2), 83–98. https://doi.org/10.1080/09658210143000281

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Gilbert, A. N., Crouch, M., & Kemp, S. E. (1998). Olfactory and Visual Mental Imagery. *Journal of Mental Imagery*. Brandon House. Retrieved from http://psycnet.apa.org/psycinfo/1998-03238-006

Gilbert, A. N., & Greenberg, M. S. (1992). Stimulus selection in the design and interpretation of olfactory studies. In M. J. Serby & K. L. Chobor (Eds.), *Science of olfaction* (pp. 309–334). New York: Springer Science & Business Media. https://doi.org/10.1007/978-1-4612-2836-3_10

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *49*(1), 585–612. https://doi.org/10.1146/annurev.psych.49.1.585

González, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuán, A., Belloch, V., & Ávila, C. (2006). Reading cinnamon activates olfactory brain regions. *NeuroImage*, *32*(2), 906–912. https://doi.org/10.1016/j.neuroimage.2006.03.037

Gorman, A. M. (1961). Recognition memory for nouns as a function of abstractness and frequency. *Journal of Experimental Psychology*, *61*(1), 23–29. https://doi.org/10.1037/h0040561

Gosmann, J., & Eliasmith, C. (2015). A Spiking Neural Model of the n-Back Task. *37th Annual Meeting of the Cognitive Science Society*, (1), 812–817. Retrieved from https://mindmodeling.org/cogsci2015/papers/0147/paper0147.pdf

Göthe, K., & Oberauer, K. (2008). The integration of familiarity and recollection information in short-term recognition: Modeling speed-accuracy trade-off functions. *Psychological Research*, *72*(3), 289–303. https://doi.org/10.1007/s00426-007-0111-9

Gottfried, J. A. (2010). Central mechanisms of odour object perception. *Nature Reviews Neuroscience*, *11*(9), 628–641. https://doi.org/10.1038/nrn2883

Gottfried, J. A., & Dolan, R. J. (2003). The nose smells what the eye sees: Crossmodal visual facilitation of human olfactory perception. *Neuron*, *39*(2), 375–386. https://doi.org/10.1016/S0896-6273(03)00392-1

Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics*. New York: Krieger.

Gruber, O., & Cramon, D. Y. Von. (2003). The functional neuroanatomy of human working memory revisited Evidence from 3-T fMRI studies using classical domain-speci c interference tasks. *Neuroimage*, *19*(3), 797–809. https://doi.org/10.1016/S1053-8119(03)00089-2

Guérard, K., & Tremblay, S. (2008). Revisiting evidence for modularity and functional equivalence across verbal and spatial domains in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 556–569. https://doi.org/10.1037/0278-7393.34.3.556

Guillery, R., & Sherman, S. (2002). Thalamic relay functions and their role in corticocortical communication: generalizations from the visual system. *Neuron*, *33*, 163–175. Retrieved from http://www.sciencedirect.com/science/article/pii/S0896627301005827

Guttentag, R. E., & Carroll, D. (1994). Identifying the basis for the word frequency effect in recognition memory. *Memory (Hove, England)*, *2*(3), 255–273. https://doi.org/10.1080/09658219408258948

Hanley, J. R., Young, A. W., & Pearson, N. A. (1991). Impairment of the visuo-spatial sketch pad. *The Quarterly Journal of Experimental Psychology Section A*, *43*(1), 101–125. https://doi.org/10.1080/14640749108401001

Harbison, J. I., Atkins, S. M., & Dougherty, M. R. (2011). N-back Training Task Performance: Analysis and Model Experiment. *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, 120–125. Retrieved from http://csjarchive.cogsci.rpi.edu/proceedings/2011/papers/0026/paper0026.pdf

Hay, D. C., Smyth, M. M., Hitch, G. J., & Horton, N. J. (2007). Serial position effects in short-term visual memory: A SIMPLE explanation? *Memory & Cognition*, *35*(1), 176–190. https://doi.org/10.3758/BF03195953

Heathcote, A., Raymond, F., & Dunn, J. C. (2006). Memory and Language Recollection and familiarity in recognition memory: Evidence from ROC curves. *Journal of Memory and Language*, *55*(4), 495–514. https://doi.org/10.1016/j.jml.2006.07.001

Hedden, T., & Park, D. C. (2003). Contributions of source and inhibitory mechanisms to age-related retroactive interference in verbal working memory. *Journal of Experimental Psychology: General*, *132*(1), 93–112. https://doi.org/10.1037/0096-3445.132.1.93

Henion, K. E. (1971). Odor pleasantness and intensity: a single dimension? *Journal of Experimental Psychology*, *90*(2), 275–279. https://doi.org/10.1037/h0031549

Herz, R. S. (2000). Verbal coding in olfactory versus nonolfactory cognition. *Memory & Cognition*, *28*(6), 957–964. https://doi.org/10.3758/BF03209343

Herz, R. S., & Cupchik, G. C. (1995). The Emotional Distinctiveness of Odor-evoked Memories. *Chemical Senses*, *20*(5), 517–528. https://doi.org/10.1093/chemse/20.5.517

Herz, R. S., & Eich, E. (1995). Commentary and envoi. In R. G. Crowder & F. R. Schab (Eds.), *Memory for Odors* (pp. 159–175). Hillsdale, New Jersey: Erlbaum.

Herz, R. S., & Engen, T. (1996). Odor memory: Review and analysis. *Psychonomic Bulletin & Review*, *3*(3), 300–13. https://doi.org/10.3758/BF03210754

Hintzman, D. L., Caulton, D. A., & Levitin, D. J. (1998). Retrieval dynamics in recognition and list discrimination: Further evidence of separate processes of familiarity and recall. *Memory & Cognition*, *26*(3), 449–462. https://doi.org/10.3758/BF03201155

Hirshman, E., & Master, S. (1997). Modeling the conscious correlates of recognition memory: Reflections on the remember-know paradigm. *Memory & Cognition*, *25*(3), 345–351. https://doi.org/10.3758/BF03211290

Hudjetz, A., & Oberauer, K. (2007). The effects of processing time and processing rate on forgetting in working memory. *Memory and Cognition*, *35*(7), 1675. https://doi.org/10.3758/BF03193501

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, *30*(6), 685–701. https://doi.org/10.1016/0749-596X(91)90032-F

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin*, *140*(2), 339–73. https://doi.org/10.1037/a0034221

Jackson, M. C., & Raymond, J. E. (2008). Familiarity enhances visual working memory for faces. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), 556–568. https://doi.org/10.1037/0096-1523.34.3.556

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513–541.

https://doi.org/10.1016/0749-596X(91)90025-F

Jacoby, L. L. (1998). Invariance in automatic influences of memory: Toward a user's guide for the process-dissociation procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(1), 3–26. https://doi.org/10.1037/0278-7393.24.1.3

Jacoby, L. L., Wahlheim, C. N., Rhodes, M. G., Daniels, J. A., & Rogers, C. S. (2010). Learning to diminish the effects of proactive interference: Reducing false memory for young and older adults. *Memory & Cognition*, *38*(6), 820–829. https://doi.org/10.1080/17439760500510569.Nice

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(19), 6829–6833. https://doi.org/10.1073/pnas.0801268105

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N -back task as a working memory measure. *Memory*, *18*(4), 394–412. https://doi.org/10.1080/09658211003702171

Jaeggi, S. M., Studer-Luethi, B., Buschkuehl, M., Su, Y. F., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning - implications for training and transfer. *Intelligence*, *38*(6), 625–635. https://doi.org/10.1016/j.intell.2010.09.001

Jalbert, A., Neath, I., & Surprenant, A. M. (2011). Does length or neighborhood size cause the word length effect? *Memory & Cognition*, *39*(7), 1198–210.

https://doi.org/10.3758/s13421-011-0094-z

Jeffreys, H. (1961). *The Theory of Probability*. OUP Oxford.

Jehl, C., Royet, J., & Holley, A. (1995). Odor discrimination and recognition memory as a function of familiarization. *Attention, Perception, & Psychophysics*, *57*(7), 1002–1011. https://doi.org/10.3758/bf03205459

Jehl, C., Royet, J., & Holley, A. (1997). Role of verbal encoding in short and long-term odor recognition. *Attention, Perception, & Psychophysics*, *59*(1), 100–110. https://doi.org/10.3758/bf03206852

Jellinek, J. S., & Köster, E. P. (1979). Perceived fragrance complexity and its relation to familiarity and pleasantness. *Journal of the Society of Cosmetic Chemists*, *30*(5), 253–262. Retrieved from http://journal.scconline.org/abstracts/cc1979/cc030n05/p00253-p00262.html

Jellinek, J. S., & Köster, E. P. (1983). Percieved fragrance complexity and its relation to familiarity and pleasantness II. *Journal of the Society of Cosmetic Chemists*, *34*(5), 253–262.

Jinks, A., & Laing, D. G. (1999). A Limit in the Processing of Components in Odour Mixtures. *Perception*, *28*(3), 395–404. https://doi.org/10.1068/p2898

Johnson, A. J., Cauchi, L., & Miles, C. (2013). Hebbian learning for olfactory sequences. *The Quarterly Journal of Experimental Psychology*, *66*(6), 1082–1089. https://doi.org/10.1080/17470218.2012.729068

Johnson, A. J., & Miles, C. (2007). Serial position functions for recognition of olfactory stimuli. *Quarterly Journal of Experimental Psychology (2006)*, *60*(10), 1347–55. https://doi.org/10.1080/17470210701515694

Johnson, A. J., & Miles, C. (2009). Single-probe serial position recall: evidence of modularity for olfactory, visual, and auditory short-term memory. *Quarterly Journal of Experimental Psychology (2006)*, *62*(2), 267–275. https://doi.org/10.1080/17470210802303750

Johnson, A. J., Volp, A., & Miles, C. (2014). Immediate recognition memory for wine. *Journal of Cognitive Psychology*, *26*(2), 127–134. https://doi.org/10.1080/20445911.2013.869225

Johnson, M. K. (1992). MEM: Mechanisms of Recollection. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn.1992.4.3.268

Johnson, M. K., Raye, C. L., Mitchell, K. J., Greene, E. J., & Anderson, A. W. (2003). fMRI evidence for an organization of prefrontal cortex by both type of process and type of information. *Cerebral Cortex*, *13*(3), 265–273. https://doi.org/10.1093/cercor/13.3.265

Johnson, M. K., Raye, C. L., Mitchell, K. J., Greene, E. J., Cunningham, W. A., & Sanislow, C. A. (2005). Using fMRI to investigate a component process of reflection: Prefrontal correlates of refreshing a just-activated representation. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(3), 339–361. https://doi.org/10.3758/CABN.5.3.339

Johnson, M. R., McCarthy, G., Muller, K. A., Brudner, S. N., & Johnson, M. K. (2015). Electrophysiological Correlates of Refreshing: Event-related Potentials Associated with Directing Reflective Attention to Face, Scene, or Word Representations. *Journal of Cognitive Neuroscience*, *27*(9), 1823–1839. https://doi.org/10.1162/jocn_a_00823

Johnson, M. R., Mitchell, K. J., Raye, C. L., D'Esposito, M., & Johnson, M. K. (2007). A brief thought can modulate activity in extrastriate visual areas: Top-down effects of refreshing just-seen visual stimuli. *NeuroImage*, *37*(1), 290–299. https://doi.org/10.1016/j.neuroimage.2007.05.017

Jones, D., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *21*(4), 1008–1018. https://doi.org/10.1037/0278-7393.21.4.1008

Jones, F. N., Roberts, K., & Holman, E. W. (1978). Similarity judgments and recognition memory for some common spices. *Perception & Psychophysics*, *24*(1), 2–6. https://doi.org/10.3758/BF03202967

Jonides, J., Badre, D., Curtis, C., Thompson-Schill, S. L., & Smith, E. E. (2002). Mechanisms of Conflict Resolution in Prefrontal Cortex. *Principles of Frontal Lobe Function*. https://doi.org/10.1093/acprof:oso/9780195134971.003.0015

Jonides, J., & Nee, D. E. (2006). Brain Mechanisms of Proactive Interference in Working Memory. *Neuroscience*, *139*(1), 181–193. https://doi.org/10.1016/j.neuroscience.2005.06.042

Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., & Koeppe, R. A. (1997). Verbal Working Memory Load Affects Regional Brain

Activation as Measured by PET. *Journal of Cognitive Neuroscience*, *9*(4), 462–475. https://doi.org/10.1162/jocn.1997.9.4.462

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proc Natl Acad Sci U S A*, *95*(14), 8410–8413. https://doi.org/10.1073/pnas.95.14.8410

Jönsson, F. U., Møller, P., & Olsson, M. J. (2011). Olfactory working memory: effects of verbalization on the 2-back task. *Memory & Cognition*, *39*(6), 1023–1032. https://doi.org/10.3758/s13421-011-0080-5

Jönsson, F. U., & Olsson, M. J. (2003). Olfactory metacognition. *Chemical Senses*, *28*(7), 651–658. https://doi.org/10.1093/chemse/bjg058

Jönsson, F. U., & Olsson, M. J. (2012). Knowing what we smell. In G. M. Zucco, R. S. Herz, & B. Schaal (Eds.), *Olfactory Cognition: Advances in Consciousness Research* (pp. 115–135). Amsterdam / Philadelphia: John Benjamins Pub. Co.

Jönsson, F. U., Tchekhova, A., Lönner, P., & Olsson, M. J. (2005). A metamemory perspective on odor naming and identification. *Chemical Senses*, *30*(4), 353–365. https://doi.org/10.1093/chemse/bji030

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*. https://doi.org/10.1007/s11065-007-9040-z

Juvina, I., & Taatgen, N. A. (2007). Modeling control strategies in the N-Back task. *8th International Conference on Cognitive Modeling*. Retrieved from http://www.academia.edu/download/30697065/juvina___taatgen.pdf

Kadohisa, M., & Wilson, D. A. (2006). Olfactory Cortical Adaptation Facilitates Detection of Odors Against Background. *Journal of Neurophysiology*, *95*(3), 1888–1896. https://doi.org/10.1152/jn.00812.2005

Kaeppler, K., & Mueller, F. (2013). Odor classification: A review of factors influencing perception-based odor arrangements. *Chemical Senses*, *38*(3), 189–209. https://doi.org/10.1093/chemse/bjs141

Kan, I. P., Barsalou, L. W., Olseth Solomon, K., Minor, J. K., & Thompson-Schill, S. L. (2003). ROLE OF MENTAL IMAGERY IN A PROPERTY VERIFICATION TASK: FMRI EVIDENCE FOR PERCEPTUAL REPRESENTATIONS OF CONCEPTUAL KNOWLEDGE. *Cognitive Neuropsychology*, *20*(3–6), 525–540. https://doi.org/10.1080/02643290244000257

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. (2007). Working memory, attention control, and the N-back task: a question of construct validity. *J Exp Psychol Learn Mem Cogn*, *33*(3), 615–622. https://doi.org/2007-06096-010 [pii]\r10.1037/0278-7393.33.3.615

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The Generality of Working Memory Capacity: A Latent-Variable Approach to Verbal and Visuospatial Memory Span and Reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189–217. https://doi.org/10.1037/0096-3445.133.2.189

Kärnekull, S. C., Jönsson, F. U., Willander, J., Sikström, S., & Larsson, M. (2015).

Long-term memory for odors: Influences of familiarity and identification across 64 days. *Chemical Senses*, *40*(4), 259–267. https://doi.org/10.1093/chemse/bjv003

Katotomichelakis, M., Balatsouras, D., Tripsianis, G., Davris, S., Maroudias, N., Danielides, V., & Simopoulos, C. (2007). The effect of smoking on the olfactory function. *Rhinology*, *45*(4), 273–280.

Keller, A. (2011). Attention and olfactory consciousness. *Frontiers in Psychology*. Frontiers. https://doi.org/10.3389/fpsyg.2011.00380

Kerr, J. R., Avons, S. E., & Ward, G. (1999). The effect of retention interval on serial position curves for item recognition of visual patterns and faces. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1475–1494. https://doi.org/10.1037//0278-7393.25.6.1475

Knops, A., Nuerk, H. C., Fimm, B., Vohn, R., & Willmes, K. (2006). A special role for numbers in working memory? An fMRI study. *NeuroImage*, *29*(1), 1–14. https://doi.org/10.1016/j.neuroimage.2005.07.009

Knudsen, E. I. (2007). Fundamental components of attention. *Annual Review of Neuroscience*, *30*, 57–78. https://doi.org/10.1146/annurev.neuro.30.051606.094256

Koen, J. D., & Yonelinas, A. P. (2016). Recollection, not familiarity, decreases in healthy ageing: Converging evidence from four estimation methods. *Memory*, *24*(1), 75–88. https://doi.org/10.1080/09658211.2014.985590

Kosslyn, S. M. (2003). Understanding the mind's eye...and nose. *Nature Neuroscience*, *6*(11), 1124. https://doi.org/10.1038/nn1103-1124

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural Foundations of Imagery. *Nature Reviews Neuroscience*, *2*(9), 635–642. https://doi.org/10.1038/35090055

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin*, *129*(5), 723–746. https://doi.org/10.1037/0033-2909.129.5.723

Köster, E. P. (2005). Does olfactory memory depend on remembering odors? In *Chemical Senses* (Vol. 30, pp. i236–i237). Oxford University Press. https://doi.org/10.1093/chemse/bjh201

Köster, E. P., Degel, J., & Piper, D. (2002). Proactive and retroactive interference in implicit odor memory. *Chemical Senses*, *27*(3), 191–206. https://doi.org/10.1093/chemse/27.3.191

Köster, E. P., Van Der Stelt, O., Nixdorf, R. R., Linschoten, M. R. I., De Wijk, R. A., & Mojet, J. (2014). Olfactory imagination and odor processing: Three same-different experiments. *Chemosensory Perception*, *7*(2), 68–84. https://doi.org/10.1007/s12078-014-9165-4

Kouider, S., de Gardelle, V., Sackur, J., & Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences*, *14*(7), 301–307. https://doi.org/10.1016/j.tics.2010.04.006

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*(4), 389–433. https://doi.org/10.1016/S0160-2896(05)80012-1

Langley, M. M., Cleary, A. M., Kostic, B. N., & Woods, J. A. (2008). Picture recognition without picture identification: A method for assessing the role of perceptual information in familiarity-based picture recognition. *Acta Psychologica*, *127*(1), 103–113. https://doi.org/10.1016/j.actpsy.2007.03.001

Larsen, J. D., & Baddeley, A. D. (2003). Disruption of verbal STM by irrelevant speech, articulatory suppression, and manual tapping: Do they have a common source? *The Quarterly Journal of Experimental Psychology Section A*, *56*(8), 1249–1268. https://doi.org/10.1080/02724980244000765

Larsson, M., Lövdén, M., & Nilsson, L. G. (2003). Sex differences in recollective experience for olfactory and verbal information. *Acta Psychologica*, *112*(1), 89–103. https://doi.org/10.1016/S0001-6918(02)00092-6

Larsson, M., Nilsson, L. G., Olofsson, J. K., & Nordin, S. (2004). Demographic and Cognitive Predictors of Cued Odor Identification: Evidence from a Population-based Study. *Chemical Senses*, *29*(6), 547–554. https://doi.org/10.1093/chemse/bjh059

Larsson, M., Öberg, C., & Bäckman, L. (2006). Recollective experience in odor recognition: Influences of adult age and familiarity. *Psychological Research*, *70*(1), 68–75. https://doi.org/10.1007/s00426-004-0190-9

Lascano, A. M., Hummel, T., Lacroix, J. S., Landis, B. N., & Michel, C. M. (2010). Spatio-temporal dynamics of olfactory processing in the human brain: An event-related source imaging study. *Neuroscience*, *167*(3), 700–708. https://doi.org/10.1016/j.neuroscience.2010.02.013

Lawless, H. T., & Cain, W. S. (1975). Recognition Memory for Odors. *Chemical Senses*, *1*(3), 331–337. https://doi.org/10.1093/chemse/1.3.331

Lawless, H. T., & Engen, T. (1977). Associations to odors: interference, mnemonics, and verbal labeling. *Journal of Experimental Psychology: Human Learning and Memory*, *3*(1), 52–59. https://doi.org/10.1037/0278-7393.3.1.52

Lehto, J. (1996). Are executive function tests dependent on working memory capacity? *Quarterly Journal of Experimental Psychology*, *49A*(1), 29–50. https://doi.org/10.1080/713755616

Leung, H.-C., & Zhang, J. X. (2004). MUY BUENO Interference resolution in spatial working memory. *NeuroImage*, *23*(3), 1013–9. https://doi.org/10.1016/j.neuroimage.2004.07.053

Levy, D., Hopkins, R., & Squire, L. (2004). Impaired odor recognition memory in patients with hippocampal lesions. *Learning & Memory*, *11*, 794–796. Retrieved from http://learnmem.cshlp.org/content/11/6/794.short

Li, W., Luxenberg, E., Parrish, T., & Gottfried, J. A. (2006). Learning to smell the roses: experience-dependent neural plasticity in human piriform and orbitofrontal cortices. *Neuron*, *52*(6), 1097–108. https://doi.org/10.1016/j.neuron.2006.10.026

Loaiza, V. M., Rhodes, M. G., Camos, V., & McCabe, D. P. (2015). Using the process dissociation procedure to estimate recollection and familiarity in working memory: An experimental and individual differences investigation. *Journal of Cognitive Psychology*, *27*(7), 844–854. https://doi.org/10.1080/20445911.2015.1033422

Logie, R. H. (1995). *Visuo-spatial Working Memory*. Psychology Press. https://doi.org/10.1002/acp.746

Logie, R. H. (2011). The Functional Organization and Capacity Limits of Working Memory. *Current Directions in Psychological Science*, *20*(4), 240–245. https://doi.org/10.1177/0963721411415340

Logie, R. H., Della Sala, S., Wynn, V., Baddeley, A. D., & Sala, S. Della. (2000). Visual similarity effects in immediate verbal serial recall. *The Quarterly Journal of Experimental Psychology Section A*, *53*(3), 626–46. https://doi.org/10.1080/713755916

Logie, R. H., Saito, S., Morita, A., Varma, S., & Norris, D. (2016). Recalling visual serial order for verbal sequences. *Memory & Cognition*, *44*(4), 590–607. https://doi.org/10.3758/s13421-015-0580-9

Logie, R. H., & Salway, A. F. S. (1990). Working memory and modes of thinking: A secondary task approach. In *Lines of thinking: Reflections on the psychology of thought, Vol. 2: Skills, emotion, creative processes, individual differences and teaching thinking.* (pp. 99–113). John Wiley & Sons. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1990-97220-007&site=ehost-live

Longoni, A. M., Richardson, J. T., & Aiello, A. (1993). Articulatory rehearsal and phonological storage in working memory. *Memory & Cognition*, *21*(1), 11–22. https://doi.org/10.3758/BF03211160

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., &

Wagenmakers, E. J. (2015). JASP (Version 0.7) [Computer Software]. Retrieved from https://jasp-stats.org

Lückmann, H. C., Jacobs, H. I. L., & Sack, A. T. (2014). The cross-functional role of frontoparietal regions in cognition: internal attention as the overarching mechanism. *Progress in Neurobiology*, *116*, 66–86. https://doi.org/10.1016/j.pneurobio.2014.02.002

Lyman, B. J., & McDaniel, M. A. (1986). Effects of encoding strategy on long-term memory for odours. *The Quarterly Journal of Experimental Psychology Section A*, *38*(4), 753–765. https://doi.org/10.1080/14640748608401624

Lyman, B. J., & McDaniel, M. A. (1990). Memory for odors and odor names: Modalities of elaboration and imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 656. https://doi.org/10.1037/0278-7393.16.4.656

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin*, *98*(1), 185–99. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4034817

Mahmut, M. K., & Stevenson, R. J. (2015). Failure to Obtain Reinstatement of an Olfactory Representation. *Cognitive Science*, *39*(8), 1940–1949. https://doi.org/10.1111/cogs.12222

Mainland, J., & Sobel, N. (2006, December 14). The sniff is part of the olfactory percept. *Chemical Senses*. Oxford University Press.

https://doi.org/10.1093/chemse/bjj012

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*(3), 252–271. https://doi.org/10.1037/0033-295X.87.3.252

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–24. https://doi.org/10.3758/s13428-011-0168-7

Mccabe, D. P., Roediger III, H. L., Mcdaniel, M. a, Balota, D. A., & Hambrick, D. Z. (2010). The Relationship Between Working Memory Capacity and Executive Functioning: Evidence for a Common Executive Attention Construct. *Neuropsychology*, *24*(2), 222–243. https://doi.org/10.1037/a0017619

McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *27*(3), 817–835. https://doi.org/10.1037/0278-7393.27.3.817

McElree, B., & Dosher, B. A. (1989). Serial Position and Set Size in Short-Term Memory: The Time Course of Recognition. *Journal of Experimental Psychology: General*, *118*(4), 346–373.

Mecklinger, A., Weber, K., Gunter, T. C., & Engle, R. W. (2003). Dissociable brain mechanisms for inhibitory control: effects of interference content and working memory capacity. *Cognitive Brain Research*, *18*(1), 26–38. https://doi.org/http://dx.doi.org/10.1016/j.cogbrainres.2003.08.008

Meiser, T., & Klauer, K. C. (1999). Working memory and changing-state hypothesis.

Journal of Experimental Psychology: Learning, Memory, and Cognition, *25*(5), 1272–1299. https://doi.org/10.1037/0278-7393.25.5.1272

Melcher, J. M., & Schooler, J. W. (1996). The Misremembrance of Wines Past: Verbal and Perceptual Expertise Differentially Mediate Verbal Overshadowing of Taste Memory. *Journal of Memory and Language*, *35*(2), 231–245. https://doi.org/10.1006/jmla.1996.0013

Miles, C., & Hodder, K. (2005). Serial position effects in recognition memory for odors: a reexamination. *Memory & Cognition*, *33*(7), 1303–14. https://doi.org/10.1037//0278-7393.26.2.411

Miles, C., & Jenkins, R. (2000). Recency and suffix effects with immediate recall of olfactory stimuli. *Memory*, *8*(3), 195–206. https://doi.org/10.1080/096582100387605

Milner, B., Corkin, S., & Teuber, H. (1968). Further analysis of the hippocampal amnesic syndrome: 14-year follow-up study of H.M. *Neuropsychologia*, *6*(3), 215–234. https://doi.org/10.1016/0028-3932(68)90021-3

Mingo, S. A., & Stevenson, R. J. (2007). Phenomenological differences between familiar and unfamiliar odours. *Perception*, *36*(6), 931–947. https://doi.org/10.1068/p5624

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Miyake, A., & Shah, P. (1999). Toward unified theories of working memory: Emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and exec- utive control* (pp. 442–481).

Moberg, P. J., Kamath, V., Marchetto, D. M., Calkins, M. E., Doty, R. L., Hahn, C. G., … Turetsky, B. I. (2014). Meta-analysis of olfactory function in schizophrenia, first-degree family members, and youths at-risk for psychosis. *Schizophrenia Bulletin*, *40*(1), 50–59. https://doi.org/10.1093/schbul/sbt049

Møller, P., Wulff, C., & Köster, E. P. (2004). Do age differences in odour memory depend on differences in verbal memory? *Neuroreport*, *15*(5), 915–917. https://doi.org/10.1097/01.wnr.0000120561.31269.ef

Monsell, S. (1978). Recency, immediate recognition memory, and reaction time. *Cognitive Psychology*, *10*(4), 465–501. https://doi.org/10.1016/0010-0285(78)90008-7

Morey, C. C., Morey, R. D., van der Reijden, M., & Holweg, M. (2013). Asymmetric cross-domain interference between two working memory tasks: Implications for models of working memory. *Journal of Memory and Language*, *69*(3), 324–348. https://doi.org/10.1016/j.jml.2013.04.004

Morey, R. D., & Rouder, J. N. (2015). BayesFactor (Version 0.9.2) [Computer Software].

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables.

Quarterly Journal of Experimental Psychology (2006), *50*(3), 528–559. https://doi.org/10.1080/027249897392017

Mueller, S. T., Seymour, T. L., Kieras, D. E., & Meyer, D. E. (2003). Theoretical implications of articulatory duration, phonological similarity, and phonological complexity in verbal working memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *29*(6), 1353–80. https://doi.org/10.1037/0278-7393.29.6.1353

Murdock, B. B. (1961). The retention of individual items. *Journal of Experimental Psychology*, *62*(6), 618–625. https://doi.org/10.1037/h0043657

Murphy, C., Cain, W. S., Gilmore, M. M., & Skinner, R. B. (1991). Sensory and semantic factors in recognition memory for odors and graphic stimuli: elderly versus young persons. *The American Journal of Psychology*, *104*(2), 161–192. https://doi.org/10.2307/1423153

Murre, J. M. J., Dros, J., Gais, S., Born, J., & Dey, M. (2015). Replication and Analysis of Ebbinghaus' Forgetting Curve. *PLOS ONE*, *10*(7), e0120644. https://doi.org/10.1371/journal.pone.0120644

Nguyen, L. A., Ober, B. A., & Shenaut, G. K. (2012). Odor recognition memory: Two encoding trials are better than one. *Chemical Senses*, *37*(8), 745–754. https://doi.org/10.1093/chemse/bjs060

Nijboer, M., Borst, J., van Rijn, H., & Taatgen, N. (2016). Contrasting single and multi-component working-memory systems in dual tasking. *Cognitive Psychology*, *86*, 1–26. https://doi.org/10.1016/j.cogpsych.2016.01.003

Nystrom, L. E., Braver, T. S., Sabb, F. W., Delgado, M. R., Noll, D. C., & Cohen, J. D. (2000). Working Memory for Letters, Shapes, and Locations: fMRI Evidence against Stimulus-Based Regional Organization in Human Prefrontal Cortex. *NeuroImage*, *11*(5 Pt 1), 424–46. https://doi.org/10.1006/nimg.2000.0572

Oberauer, K. (2005). Binding and inhibition in working memory: individual and age differences in short-term recognition. *Journal of Experimental Psychology. General*, *134*(3), 368–387. https://doi.org/10.1037/0096-3445.134.3.368

Oberauer, K. (2009). *Design for a working memory*. *Psychology of Learning and Motivation* (Vol. 51). https://doi.org/10.1016/S0079-7421(09)51002-X

Oberauer, K., Süß, H. M., Schulze, R., Wilhelm, O., & Wittmann, W. W. (2000). Working memory capacity — facets of a cognitive ability construct. *Personality and Individual Differences*, *29*(6), 1017–1045. https://doi.org/10.1016/S0191-8869(99)00251-2

Öberg, C., Larsson, M., & Bäckman, L. (2002). Differential sex effects in olfactory functioning: The role of verbal processing. *Journal of the International Neuropsychological Society*, *8*(5), 691–698. https://doi.org/10.1017/S1355617702801424

Olsson, H., & Poom, L. (2005). Visual memory needs categories. *Proceedings of the National Academy of Sciences*, *102*(24), 8776–8780. https://doi.org/10.1073/pnas.0500810102

Olsson, M. J., Lundgren, E. B., Soares, S. C., & Johansson, M. (2009). Odor memory performance and memory awareness: A comparison to word memory across orienting tasks and retention intervals. *Chemosensory Perception*, *2*(3), 161–171. https://doi.org/10.1007/s12078-009-9051-7

Ongür, D., & Price, J. L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral Cortex*, *10*(3), 206–219. https://doi.org/10.1093/cercor/10.3.206

Owen, A. M. (1997). The functional organization of working memory processes within human lateral frontal cortex: the contribution of functional neuroimaging. *The European Journal of Neuroscience*, *9*(7), 1329–39. https://doi.org/10.1111/j.1460-9568.1997.tb01487.x

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, *25*(1), 46–59. https://doi.org/10.1002/hbm.20131

Owen, A. M., Morris, R., Sahakian, B., & Polkey, C. (1996). Double dissociations of memory and executive functions in working memory tasks following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Brain*, *119*, 1597–1615. Retrieved from https://academic.oup.com/brain/article-abstract/119/5/1597/369214

Paivio, A. (1990). *Mental representations. A Dual Coding Approach*. Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal Of Psychology*, *45*(3), 255–287. https://doi.org/10.1037/h0084295

Peynircioğlu, Z. F. (1990). A feeling-of-recognition without identification. *Journal of Memory and Language*, *29*(4), 493–500. https://doi.org/10.1016/0749-596X(90)90068-B

Pickering, S. J. (2001). The development of visuo-spatial working memory. *Memory*, *9*(4–6), 423–432. https://doi.org/10.1080/09658210143000182

Pinault, D. (2004, August). The thalamic reticular nucleus: Structure, function and concept. *Brain Research Reviews*. https://doi.org/10.1016/j.brainresrev.2004.04.008

Plailly, J., Delon-Martin, C., & Royet, J. P. (2012). Experience induces functional reorganization in brain regions involved in odor imagery in perfumers. *Human Brain Mapping*, *33*(1), 224–234. https://doi.org/10.1002/hbm.21207

Portas, C. M., Rees, G., Howseman, A. M., Josephs, O., Turner, R., & Frith, C. D. (1998). A specific role for the thalamus in mediating the interaction of attention and arousal in humans. *J Neurosci*, *18*(21), 8979–8989. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9787003

Postle, B. R., & Brush, L. N. (2004). The neural bases of the effects of item-nonspecific proactive interference in working memory. *Cogn Affect Behav Neurosci*, *4*(3), 379–392. https://doi.org/10.3758/CABN.4.3.379

Postle, B. R., Brush, L. N., & Nick, A. M. (2004). Prefrontal cortex and the mediation of proactive interference in working memory. *Cognitive, Affective & Behavioral Neuroscience*, *4*(4), 600–608. https://doi.org/10.3758/CABN.4.4.600

Postman, L., & Underwood, B. J. (1973). Critical issues in interference theory. *Memory & Cognition*, *1*(1), 19–40. https://doi.org/10.3758/BF03198064

Prabhakaran, R., & Thompson-Schill, S. L. (2011). The Price of Fame: The Impact of Stimulus Familiarity on Proactive Interference Resolution. *Journal of Cognitive Neuroscience*, *23*(4), 816–831. https://doi.org/10.1162/jocn.2010.21501

Prull, M. W., Dawes, L. L., Martin 3rd, A. M., Rosenberg, H. F., & Light, L. L. (2006). Recollection and familiarity in recognition memory: adult age differences and neuropsychological test correlates. *Psychol Aging*, *21*(1), 107–118. https://doi.org/10.1037/0882-7974.21.1.107

Purdon, S. E., Klein, S., & Flor-Henry, P. (2001). Menstrual effects on asymmetrical olfactory acuity. *Journal of the International Neuropsychological Society : JINS*, *7*(6), 703–9. https://doi.org/10.1017/S1355617701766064

Pylyshyn, Z. (2003). Return of the mental image: are there really pictures in the brain? *Trends in Cognitive Sciences*, *7*(3), 113–118. Retrieved from http://www.sciencedirect.com/science/article/pii/S1364661303000032

Rabin, M. D. (1988). Experience facilitates olfactory quality discrimination. *Perception & Psychophysics*, *44*(6), 532–540. https://doi.org/10.3758/BF03207487

Rabin, M. D., & Cain, W. S. (1984). Odor recognition: Familiarity, identifiability, and encoding consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(2), 316–325. https://doi.org/http://dx.doi.org/10.1037/0278-7393.10.2.316

Ragland, J. D., Turetsky, B. I., Gur, R. C., Gunning-Dixon, F., Turner, T., Schroeder, L., … Gur, R. E. (2002). Working memory for complex figures: an fMRI comparison of letter and fractal n-back tasks. *Neuropsychology*, *16*(3), 370–379. https://doi.org/10.1037/0894-4105.16.3.370

Rajaram, S. (1993). Remembering and knowing: two means of access to the personal past. *Memory & Cognition*, *21*(1), 89–102. https://doi.org/10.3758/BF03211168

Rajaram, S. (1998). The effects of conceptual salience and perceptual distinctiveness on conscious recollection. *Psychonomic Bulletin & Review*, *5*(1), 71–78. https://doi.org/10.3758/BF03209458

Ralph, J. (2014). *Statistical manipulation and control strategies of the n-back task*.

Ranganath, C., & D'Esposito, M. (2005, April). Directing the mind's eye: Prefrontal, inferior and medial temporal mechanisms for visual working memory. *Current Opinion in Neurobiology*. https://doi.org/10.1016/j.conb.2005.03.017

Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: A minimal executive function. In *Cortex* (Vol. 43, pp. 135–145). https://doi.org/10.1016/S0010-9452(08)70451-9

Raye, C. L., Johnson, M. K., Mitchell, K. J., Reeder, J. A., & Greene, E. J. (2002). Neuroimaging a single thought: dorsolateral PFC activity associated with refreshing just-activated information. *NeuroImage*, *15*(2), 447–53. https://doi.org/10.1006/nimg.2001.0983

Reder, L. M., Liu, X. L., Keinath, A., & Popov, V. (2015). Building knowledge requires bricks, not sand: The critical role of familiar constituents in learning. *Psychonomic Bulletin & Review*, 1–7. https://doi.org/10.3758/s13423-015-0889-1

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember–know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 294–320. https://doi.org/http://dx.doi.org/10.1037/0278-7393.26.2.294

Redick, T. S., & Lindsey, D. R. B. (2013). Complex span and n-back measures of working memory: a meta-analysis. *Psychonomic Bulletin & Review*, *20*(6), 1102–13. https://doi.org/10.3758/s13423-013-0453-9

Reed, P. (2000). Serial Position Effects in Recognition Memory for Odors. *Journal of Experimental Psychology: Learning Memory and Cognition*, *26*(2), 411–422. Retrieved from https://www.scopus.com/inward/record.url?eid=2-s2.0-0034154649&partnerID=40&md5=9e4b81a094221b6587a9ae6e8da0efbb

Richardson, J. T. E. (1975). Concreteness and imageability. *Quarterly Journal of Experimental Psychology*, *27*(February 2015), 235–249. https://doi.org/10.1080/14640747508400483

Richardson, J. T. E., & Baddeley, A. D. (1975). The effect of articulatory suppression in free recall. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 623–629. https://doi.org/10.1016/S0022-5371(75)80049-1

Rinck, F., Rouby, C., & Bensafi, M. (2009). Which format for odor images? *Chemical*

Senses, *34*(1), 11–13. https://doi.org/10.1093/chemse/bjn060

Roediger III, H. L., & DeSoto, K. A. (2014). Confidence and memory: assessing positive and negative correlations. *Memory (Hove, England)*, *22*(1), 76–91. https://doi.org/10.1080/09658211.2013.795974

Rolls, E. T., Grabenhorst, F., & Margot, C. (2008). Selective attention to affective value alters how the brain processes olfactory stimuli. *Journal of Cognitive Neuroscience*, *20*(10), 1815–1826. Retrieved from http://www.mitpressjournals.org/doi/abs/10.1162/jocn.2008.20128

Rolls, E. T., Kringelbach, M. L., & De Araujo, I. E. T. (2003). Different representations of pleasant and unpleasant odours in the human brain. *European Journal of Neuroscience*, *18*(3), 695–703. https://doi.org/10.1046/j.1460-9568.2003.02779.x

Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (2004). Sum-Difference Theory of Remembering and Knowing: A Two-Dimensional Signal-Detection Model. *Psychological Review*, *111*(3), 588–616. https://doi.org/10.1037/0033-295X.111.3.588

Rotello, C. M., Macmillan, N. A., Reeder, J. A., & Wong, M. (2005). The remember response: subject to bias, graded, and not a process-pure indicator of recollection. *Psychonomic Bulletin & Review*, *12*(5), 865–73. https://doi.org/10.3758/BF03196778

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Royet, J. P., Delon-Martin, C., & Plailly, J. (2013). Odor mental imagery in non-experts in odors: a paradox? *Front Hum Neurosci*, *7*(March), 87. https://doi.org/10.3389/fnhum.2013.00087

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, *89*(1), 63–77. https://doi.org/10.1037/h0031185

Saito, S., Logie, R. H., Morita, A., & Law, A. (2008). Visual and phonological similarity effects in verbal immediate serial recall: A test with kanji materials. *Journal of Memory and Language*, *59*(1), 1–17. https://doi.org/10.1016/j.jml.2008.01.004

Salamé, P., & Baddeley, A. D. (1986). Phonological factors in STM: Similarity and the unattended speech effect. *Bulletin of the Psychonomic Society*, *24*(4), 263–265. https://doi.org/10.3758/BF03330135

Sauvage, M. M., Beer, Z., & Eichenbaum, H. (2010). Recognition memory: Adding a response deadline eliminates recollection but spares familiarity. *Learning & Memory*, *17*(2), 104–108. https://doi.org/10.1101/lm.1647710

Sauvage, M. M., Fortin, N. J., Owens, C. B., Yonelinas, A. P., & Eichenbaum, H. (2008). Recognition memory: opposite effects of hippocampal damage on recollection and familiarity. *Nature Neuroscience*, *11*(1), 16–18. https://doi.org/10.1038/nn2016

Savic, I., & Berglund, H. (2000). Right-nostril dominance in discrimination of unfamiliar, but not familiar, odours. *Chemical Senses*, *25*(5), 517–523. https://doi.org/10.1093/chemse/25.5.517

Schab, F. R. (1991). Odor memory: Taking stock. *Psychological Bulletin*, *109*(2), 242–251. https://doi.org/10.1037/0033-2909.109.2.242

Schab, F. R., & Crowder, R. G. (1995). *Memory for odors / edited by Frank R. Schab, Robert G. Crowder*. Mahwah, N.J. : L. Erlbaum Associates, c1995. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=edshlc&AN=edshlc.005649614-1&site=eds-live&scope=site

Schiffman, S. S. (1974). Physicochemical correlates of olfactory quality. A series of physicochemical variables are weighted mathematically to predict olfactory quality. *Science*, *185*(4146), 112–117. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=edselc&AN=edselc.2-52.0-0016184331&site=eds-live&scope=site

Schmiedek, F., Hildebrandt, A., Lövdén, M., Lindenberger, U., & Wilhelm, O. (2009). Complex span versus updating tasks of working memory: the gap is not that deep. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(4), 1089–1096. https://doi.org/10.1037/a0015730

Schmiedek, F., Li, S.-C., & Lindenberger, U. (2009). Interference and facilitation in spatial working memory: age-associated differences in lure effects in the n-back paradigm. *Psychology and Aging*, *24*(1), 203–210. https://doi.org/10.1037/a0014685

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2014). A task is a task is a task: Putting complex span, n-back, and other working memory indicators in psychometric context. *Frontiers in Psychology*, *5*(DEC).

https://doi.org/10.3389/fpsyg.2014.01475

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal Overshadowing of Visual Memories: Some Things Are better Left Unsaid. *Cognitive Psychology*, *22*, 36–71. Retrieved from http://ac.els-cdn.com/001002859090003M/1-s2.0-001002859090003M-main.pdf?_tid=0fc6fca2-3cbf-11e7-804f-00000aacb35e&acdnat=1495217876_2326fcd21b6290e5d7c0b540c48f1cd3

Sezille, C., Fournel, A., Rouby, C., Rinck, F., & Bensafi, M. (2014). Hedonic appreciation and verbal description of pleasant and unpleasant odors in untrained, trainee cooks, flavorists, and perfumers. *Frontiers in Psychology*, *5*(JAN), 12. https://doi.org/10.3389/fpsyg.2014.00012

Shelton, J. T., Elliott, E. M., Hill, B. D., Calamia, M. R., & Gouvier, W. (2009). A comparison of laboratory and clinical working memory tests and their prediction of fluid intelligence. *Intelligence*, *37*(3), 283–293. https://doi.org/10.1016/j.intell.2008.11.005.A

Shelton, J. T., Elliott, E. M., Matthews, R. A., Hill, B. D., & Gouvier, W. (2010). The relationships of working memory, secondary memory, and general fluid intelligence: working memory is special. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(3), 813–20. https://doi.org/10.1037/a0019046

Shelton, J. T., Metzger, R. L., & Elliott, E. M. (2007). A group-administered lag task as a measure of working memory. *Behavior Research Methods*, *39*(3), 482–493. https://doi.org/10.3758/BF03193017

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, *6*(1), 156–163. https://doi.org/10.1016/S0022-5371(67)80067-7

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science (New York, N.Y.)*, *171*(972), 701–703. https://doi.org/10.1126/science.171.3972.701

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences*, *12*(April), 182–186. https://doi.org/10.1016/j.tics.2008.02.003

Simmons, M. R. (2000). *The central executive and working memory: A dual-task investigation of the n-back task*. *ProQuest Dissertations and Theses*. ProQuest Information & Learning, US. Retrieved from http://ua.lm.worldcat.org/?genre=article&sid=ProQ:&atitle=&title=The+central+executive+and+working+memory%3A+A+dual-task+investigation+of+the+n-back+task&issn=&date=2000-01-01&volume=&issue=&spage=&author=Simmons%2C+Michelle+R

Sinding, C., Coureaud, G., Bervialle, B., Martin, C., Schaal, B., & Thomas-Danguin, T. (2015). Experience shapes our odor perception but depends on the initial perceptual processing of the stimulus. *Attention, Perception, & Psychophysics*, *77*(5), 1794–1806. https://doi.org/10.3758/s13414-015-0883-8

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*, 1657–1661. Retrieved from http://science.sciencemag.org/content/283/5408/1657.short

Smith, E. E., Jonides, J., Koeppe, R. a., Awh, E., Schumacher, E. H., & Minoshima, S. (1995). Spatial versus Object Working Memory: PET Investigations. *Journal of Cognitive Neuroscience*, *7*, 337–356. https://doi.org/10.1162/jocn.1995.7.3.337

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, *6*(2), 174–215. https://doi.org/10.1037/0278-7393.6.2.174

Souza, A. S., Rerko, L., & Oberauer, K. (2015). Refreshing memory traces: Thinking of an item improves retrieval from visual working memory. *Annals of the New York Academy of Sciences*, *1339*(1), 20–31. https://doi.org/10.1111/nyas.12603

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, {&} Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Stevenson, R. J. (2001). Perceptual learning with odors: implications for psychological accounts of odor quality perception. *Psychonomic Bulletin & Review*, *8*(4), 708–712. https://doi.org/10.3758/BF03196207

Stevenson, R. J. (2009). Phenomenal and access consciousness in olfaction. *Consciousness and Cognition*, *18*(4), 1004–1017. https://doi.org/10.1016/j.concog.2009.09.005

Stevenson, R. J. (2010). An initial evaluation of the functions of human olfaction. *Chemical Senses*. https://doi.org/10.1093/chemse/bjp083

Stevenson, R. J. (2012). Olfactory Perception. In *Olfactory Cognition: Advances in Consciousness Research* (pp. 73–91).

Stevenson, R. J., & Attuquayefio, T. (2013). Human olfactory consciousness and cognition: Its unusual features may not result from unusual functions but from limited neocortical processing resources. *Frontiers in Psychology*, *4*(NOV), 819. https://doi.org/10.3389/fpsyg.2013.00819

Stevenson, R. J., & Boakes, R. A. (2003). A Mnemonic Theory of Odor Perception. *Psychological Review*, *110*(2), 340–364. https://doi.org/10.1037/0033-295X.110.2.340

Stevenson, R. J., Boakes, R. A., & Wilson, J. P. (2000). Resistance to extinction of conditioned odor perceptions: Evaluative conditioning is not unique. *Journal of Experimental Psychology-Learning Memory and Cognition*, *26*(2), 423–440. https://doi.org/10.1037/0278-7393.26.2.423

Stevenson, R. J., & Case, T. I. (2005). Olfactory imagery: A review. *Psychonomic Bulletin & Review*, *12*(2), 244–264. https://doi.org/10.3758/BF03196369

Stevenson, R. J., Case, T. I., & Mahmut, M. K. (2007). Difficulty in evoking odor images: the role of odor naming. *Memory & Cognition*, *35*(3), 578–589. https://doi.org/10.3758/BF03193296

Stevenson, R. J., & Mahmut, M. K. (2013a). Familiarity influences odor memory stability. *Psychonomic Bulletin & Review*, *20*(4), 754–9. https://doi.org/10.3758/s13423-013-0380-9

Stevenson, R. J., & Mahmut, M. K. (2013b). The accessibility of semantic knowledge for odours that can and cannot be named. *Q J Exp Psychol (Hove)*, *66*(7), 1414–1431. https://doi.org/10.1080/17470218.2012.753097

Stevenson, R. J., & Wilson, D. A. (2007). Odour perception: An object-recognition approach. *Perception*, *36*(12), 1821–1833. https://doi.org/10.1068/p5563

Sugiyama, H., Ayabe-Kanamura, S., & Kikuchi, T. (2006). Are olfactory images sensory in nature? *Perception*, *35*(12), 1699–1708. https://doi.org/10.1068/p5453

Sulmont, C., Issanchou, S., & Koster, E. P. (2002). Selection of Odorants for Memory Tests on the Basis of Familiarity, Perceived Complexity, Pleasantness, Similarity and Identification. *Chemical Senses*, *27*(4), 307–317. https://doi.org/10.1093/chemse/27.4.307

Süß, H. M., Oberauer, K., Wittman, W. W., Wilhelm, O., & Schulze, R. (2002). Working memory capacity explains reasoning ability—And a little bit more. *Intelligence*, *30*, 261–288. Retrieved from http://www.sciencedirect.com/science/article/pii/S0160289601001003

Szmalec, A., Verbruggen, F., Vandierendonck, A., & Kemps, E. (2011). Control of interference during working memory updating. *Journal of Experimental Psychology. Human Perception and Performance*, *37*(1), 137–151. https://doi.org/10.1037/a0020365

Tan, L., & Ward, G. (2000). A recency-based account of the primacy effect in free recall. *Journal Of Experimental Psychology-Learning Memory And Cognition*, *26*(6), 1589–1625. https://doi.org/10.1037/0278-7393.26.6.1589

Tan, L., & Ward, G. (2007). Output order in immediate serial recall. *Memory & Cognition*, *35*(5), 1093–106. https://doi.org/10.3758/BF03193481

Thomas-Danguin, T., Sinding, C., Romagny, S., El Mountassir, F., Atanasova, B., Le Berre, E., … Coureaud, G. (2014). The perception of odor objects in everyday life: a review on the processing of odor mixtures. *Frontiers in Psychology*, *5*, 504. https://doi.org/10.3389/fpsyg.2014.00504

Thorn, A. S. C., Gathercole, S. E., & Frankish, C. R. (2002). Language familiarity effects in short-term memory: The role of output delay and long-term knowledge. *The Quarterly Journal of Experimental Psychology Section A*, *55*(4), 1363–1383. https://doi.org/10.1080/02724980244000198

Tomiczek, C., & Stevenson, R. J. (2009). Olfactory Imagery and Repetition Priming: The Effect of Odor Naming and Imagery Ability. *Experimental Psychology*, *56*(6), 397–408. https://doi.org/10.1027/1618-3169.56.6.397

Tong, F. (2013, October). Imagery and visual working memory: One and the same? *Trends in Cognitive Sciences*. NIH Public Access. https://doi.org/10.1016/j.tics.2013.08.005

Tulving, E. (1985). Memory and Consciousness. *Canadian Psychology-Psychologie Canadienne*, *26*(1), 1–12. https://doi.org/10.1037/h0080017

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154. https://doi.org/10.1016/0749-596X(89)90040-5

Uebersax, J. (2015). Agreement on Interval-Level Ratings. Retrieved from http://www.john-uebersax.com/stat/agree.htm

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, *133*(6), 1038–1066. https://doi.org/10.1037/0033-2909.133.6.1038

Unsworth, N., & Spillers, G. J. (2010). Variation in working memory capacity and episodic recall: The contributions of strategic encoding and contextual retrieval. *Psychonomic Bulletin & Review*, *17*(2), 200–205. https://doi.org/10.3758/PBR.17.2.200

Valentin, D., Dacremont, C., & Cayeux, I. (2011). Does short-term odour memory increase with expertise? An experimental study with perfumers, flavourists, trained panellists and novices. *Flavour and Fragrance Journal*, *26*(6), 408–415. https://doi.org/10.1002/ffj.2069

Vallar, G., & Baddeley, A. D. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behavior*, *23*(2), 151–161. https://doi.org/10.1016/S0022-5371(84)90104-X

Vandierendonck, A. (2016). Modality independence of order coding in working memory: Evidence from cross-modal order interference at recall. *The Quarterly Journal of Experimental Psychology*, *69*(1), 161–179. https://doi.org/10.1080/17470218.2015.1032987

Verhagen, J. V, Wesson, D. W., Netoff, T. I., White, J. A., & Wachowiak, M. (2007). Sniffing controls an adaptive filter of sensory input to the olfactory bulb. *Nature Neuroscience*, *10*(5), 631–639. https://doi.org/10.1038/nn1892

Vuontela, V., Rämä, P., Raninen, A., Aronen, H. J., & Carlson, S. (1999). Selective interference reveals dissociation between memory for location and colour. *Neuroreport*, *10*(11), 2235–40. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10439440

Wahlheim, C. N., & Jacoby, L. L. (2011). Experience with proactive interference diminishes its effects: mechanisms of change. *Memory & Cognition*, *39*(2), 185–195. https://doi.org/10.3758/s13421-010-0017-4

Wais, P. E., Mickes, L., & Wixted, J. T. (2008). Remember/know judgments probe degrees of recollection. *Journal of Cognitive Neuroscience*, *20*(3), 400–405. https://doi.org/10.1162/jocn.2008.20041

Walk, H. A., & Johns, E. E. (1984). Interference and facilitation in short-term memory for odors. *Perception & Psychophysics*, *36*(6), 508–14. https://doi.org/10.3758/BF03207510

Ward, G., Avons, S. E., & Melling, L. (2005). Serial position curves in short-term memory: Functional equivalence across modalities. *Memory*, *13*(3–4), 308–317. https://doi.org/10.1080/09658210344000279

Watter, S., Geffen, G. M., & Geffen, L. B. (2001). The n-back as a dual-task: P300 morphology under divided attention. *Psychophysiology*, *38*(6), 998–1003. https://doi.org/10.1111/1469-8986.3860998

White, T. L. (1998). Olfactory memory: the long and short of it. *Chemical Senses*, *23*(4), 433–441. https://doi.org/10.1093/chemse/23.4.433

White, T. L. (2009). A second look at the structure of human olfactory memory. *Annals of the New York Academy of Sciences*, *1170*(1), 338–342. https://doi.org/10.1111/j.1749-6632.2009.03878.x

White, T. L. (2012). Attending to olfactory short-term memory. In G. M. Zucco, R. S. Herz, & B. Schaal (Eds.), *Olfactory Cognition: From perception and memory to environmental odours and neuroscience* (p. 317). Amsterdam / Philadelphia: John Benjamins Pub. Co.

White, T. L., Hornung, D. E., Kurtz, D. B., Treisman, M., & Sheehe, P. (1998). Phonological and perceptual components of short-term memory for odors. *American Journal of Psychology*, *111*(3), 411–434. https://doi.org/10.1111/j.1749-6632.1998.tb10636.x

White, T. L., Møller, P., Köster, E. P., Eichenbaum, H., & Linster, C. (2015). Olfactory Memory. In *Handbook of Olfaction and Gustation* (Vol. 1170, pp. 337–352). https://doi.org/10.1002/9781118971758.ch15

White, T. L., & Treisman, M. (1997). A comparison of the encoding of content and order in olfactory memory and in memory for visually presented verbal materials. *British Journal of Psychology (London, England : 1953)*, *88 (Pt 3)*(February 2016), 459–472. https://doi.org/10.1111/j.2044-8295.1997.tb02651.x

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology*, *4*(JUL), 1–22.

https://doi.org/10.3389/fpsyg.2013.00433

Wilson, D. A., & Stevenson, R. J. (2003a). Olfactory perceptual learning: The critical role of memory in odor discrimination. *Neuroscience and Biobehavioral Reviews*, *27*(4), 307–328. https://doi.org/10.1016/S0149-7634(03)00050-2

Wilson, D. A., & Stevenson, R. J. (2003b). The fundamental role of memory in olfactory perception. *Trends in Neurosciences*, *26*(5), 243–247. https://doi.org/10.1016/S0166-2236(03)00076-6

Wilson, D. A., & Stevenson, R. J. (2006). *Learning to smell: olfactory perception from neurobiology to behavior*. *Physiological Soc*. JHU Press.

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review*, *117*(4), 1025–1054. https://doi.org/10.1037/a0020874

Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychonomic Bulletin & Review*, *11*(4), 616–641. https://doi.org/10.3758/BF03196616

Wood, J. B., & Harkins, S. W. (1987). Effects of age, stimulus selection, and retrieval environment on odor identification. *Journal of Gerontology*, *42*(6), 584–588.

Yeshurun, Y., Dudai, Y., & Sobel, N. (2008). Working memory across nostrils. *Behavioral Neuroscience*, *122*(5), 1031–1037. https://doi.org/10.1037/a0012806

Yonelinas, A. P. (1999). The contribution of recollection and familiarity contributions to recognition and source memory judgments: A formal dual-process model and an analysis of receiver operating charateristics.pdf. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *25*(6), 1415–1434.

Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., Dobbins, I., Szymanski, M. D., Dhaliwal, H. S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Consciousness and Cognition*, *5*(4), 418–441. https://doi.org/10.1006/ccog.1996.0026

Yonelinas, A. P., & Jacoby, L. L. (1994). Dissociations of processes in recognition memory: Effects of interference and of response speed. *Canadian Journal of Experimental Psychology*, *48*(4), 516–535. https://doi.org/10.1037/1196-1961.48.4.516

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: a review. *Psychological Bulletin*, *133*(5), 800–32. https://doi.org/10.1037/0033-2909.133.5.800

Yoon, C., Feinberg, F., Luo, T., Hedden, T., Gutchess, A. H., Chen, H. Y., … C., P. D. (2004). A Cross-Culturally Standardized Set of Pictures for Younger and Older Adults: American and Chinese Norms for Name Agreement, Concept Agreement and Familiarity. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 639–649. Retrieved from http://agingmind.utdallas.edu/stimuli/culturedb/pictnorms/

Yoshida, M. (1964). Studies of psychometric classification of odors: V. *Japanese Psychological Research*, *6*(4), 145–154.

Zajonc, R. B. (1968). Attitudinal Effects of Mere Exposure. *Journal of Personality and Social Psychology*, *9*(2, Pt.2), 1–27. https://doi.org/10.1037/h0025848

Zelano, C., Montag, J., Khan, R., & Sobel, N. (2009). A specialized odor memory buffer in primary olfactory cortex. *PLoS ONE*, *4*(3), e4965–e4965. https://doi.org/10.1371/journal.pone.0004965

Zucco, G. M. (2003). Anomalies in cognition: Olfactory memory. *European Psychologist*, *8*(2), 77–86. https://doi.org/10.1027//1016-9040.8.2.77

8 Appendices

8.1 Appendix A: Normative data

Odour	Fam.	Int.	Pleas.	Irr.	Cont. Av.	Hed. Str.	Comp.	AoA	Freq.	Desc.	Verb.
Almonds	5.34	4.60	5.12	2.80	4.28	1.36	4.20	9.03	3.74	4.16	2.27
Aniseed Balls	5.88	5.40	3.98	3.54	5.28	1.50	4.08	10.02	4.22	5.04	2.61
Apples Green	4.75	4.21	5.04	2.83	4.02	1.50	3.62	10.30	3.83	3.88	2.04
Bacon	3.15	4.35	2.90	4.46	2.65	1.27	4.08	14.93	2.38	2.21	1.15
Banana	5.40	4.76	5.32	2.94	4.54	1.56	3.88	9.89	4.18	4.48	2.18
Barbecue	3.62	5.26	2.30	5.10	3.14	2.02	4.33	13.44	3.08	2.98	1.55
Basil	4.06	4.43	4.12	3.37	3.35	1.22	4.38	12.66	3.39	2.96	1.64
Beef	3.80	6.20	1.88	5.67	3.47	2.33	4.71	13.88	2.69	3.41	1.73
Biscuit	4.17	4.13	3.48	3.98	3.50	1.31	4.09	12.17	3.33	3.25	1.80
Blackberry	5.04	4.23	5.23	2.54	4.48	1.52	3.83	9.66	3.71	4.46	2.09
Blackcurrant	5.67	4.85	5.48	2.52	5.13	1.73	3.91	7.12	4.62	4.67	2.44
Blue Cheese	3.51	5.66	1.89	5.60	2.87	2.19	4.47	14.29	2.55	2.94	1.66

Brandy	4.28	4.60	3.96	3.92	3.66	1.40	3.92	14.46	3.64	3.42	1.76
Bubble Gum	5.53	5.04	5.41	2.49	5.04	1.61	4.19	8.73	4.27	4.86	2.34
Butter Cream	5.02	4.71	4.35	3.56	4.08	1.23	3.71	10.91	3.40	3.79	2.04
Buttered Popcorn	2.72	4.58	3.02	4.42	2.80	1.47	4.16	15.07	2.68	2.70	1.19
Cabbage	3.00	6.24	1.31	6.34	2.69	2.76	4.10	15.32	2.31	2.48	1.21
Candy Floss	3.43	4.04	3.65	4.00	2.82	1.04	3.88	14.11	2.90	2.54	1.38
Caramel Toffee	4.16	3.58	4.28	3.32	3.46	1.24	3.49	11.84	3.18	3.30	1.53
Cardamom	4.08	5.08	3.32	4.16	2.92	1.52	4.22	13.51	2.92	2.84	1.54
Carrot	3.55	4.69	2.88	4.45	2.71	1.49	3.90	14.57	2.67	2.71	1.29
Celery	3.26	4.70	2.80	4.86	2.96	1.60	4.24	13.78	2.78	2.80	1.59
Cereal	4.17	4.90	3.29	4.53	3.96	1.37	3.73	11.40	3.31	3.45	2.03
Cheddar Cheese	3.14	5.27	2.35	5.29	2.79	1.86	4.04	15.90	2.53	2.63	1.24
Cherry	5.58	4.76	5.28	2.78	4.45	1.80	3.73	9.31	3.88	4.38	2.28
Chewing Gum	5.85	5.02	5.11	3.04	5.11	1.53	3.75	8.80	4.81	4.70	2.64
Chicken	3.40	4.48	2.85	4.58	2.81	1.69	3.91	13.74	2.96	2.67	1.34
Chilli Pepper	3.00	4.14	3.22	4.29	2.69	0.98	4.25	15.71	2.54	2.55	1.31

Chocolate	4.67	4.41	3.80	3.37	3.94	1.35	3.98	11.12	3.82	3.71	2.06
Chocolate (Mint)	4.90	5.48	4.12	3.84	4.22	1.32	4.27	11.11	3.63	4.06	2.24
Chocolate (Orange)	4.24	5.04	4.00	3.74	3.54	1.40	4.35	11.20	3.44	3.24	1.88
Cinder Toffee	3.02	3.37	3.59	3.91	2.26	1.04	3.39	14.80	2.53	2.47	1.02
Cinnamon	4.54	5.00	4.44	3.52	3.63	1.27	4.26	12.78	3.50	3.38	1.79
Coco-mango	5.27	4.25	5.55	2.39	4.61	1.64	3.61	10.42	4.42	4.20	2.05
Coconut	4.87	3.23	4.83	2.85	4.19	1.21	3.59	10.24	3.94	3.63	2.15
Coffee	4.45	4.88	3.76	3.84	4.24	1.47	4.13	10.97	3.80	3.82	2.05
Cola	5.21	5.04	5.02	3.02	4.29	1.35	4.02	10.05	3.89	3.81	2.14
Cookie	3.94	3.94	4.33	3.45	3.12	1.14	4.19	14.28	2.94	2.88	1.35
Cookies & Cream	4.30	4.92	3.80	3.98	3.80	1.28	4.24	12.70	3.56	3.64	1.94
Coriander	3.87	4.76	3.98	3.87	2.96	1.11	3.98	12.73	2.98	2.89	1.51
Cranberry	5.59	4.71	5.58	2.59	4.73	1.83	3.88	9.25	4.55	4.49	2.29
Crusty Bread	3.84	3.76	3.57	3.76	3.00	1.00	3.69	12.68	3.58	3.24	1.38
Cucumber & Lime	5.34	4.28	5.38	2.72	4.83	1.51	3.91	11.66	4.55	4.40	2.34
Cumin	4.27	4.96	3.30	4.58	3.46	1.44	4.11	13.13	3.67	3.00	1.57

Curry	4.81	5.29	3.79	4.02	3.77	1.29	4.28	11.70	3.73	3.40	2.13
Fermented Fruit	4.18	5.24	3.62	4.14	3.58	1.38	4.49	14.16	3.20	3.42	1.77
Fish	4.74	5.78	1.66	5.66	4.74	2.50	3.69	10.50	3.68	4.58	2.57
Fruit Punch	5.45	4.18	5.29	2.49	5.06	1.53	3.77	9.46	4.63	4.63	2.45
Fruity Sweets	5.02	4.40	5.04	2.83	4.38	1.38	3.91	10.65	4.40	4.02	1.91
Garden Mint	5.84	5.08	5.30	2.32	5.60	1.66	3.77	8.08	5.24	5.22	2.62
Antiseptic	3.48	4.00	3.30	4.19	3.35	1.13	4.17	13.04	3.13	3.08	1.41
Baby powder	5.48	4.25	5.46	2.48	4.67	1.63	3.70	9.83	4.28	4.25	2.18
Beauty Soap	4.98	4.12	4.71	2.96	4.71	1.33	3.88	9.05	4.18	4.31	2.21
Black Pepper	4.88	5.04	4.24	3.66	4.00	1.28	4.18	10.97	3.84	3.68	1.77
Boiler Room	3.04	5.86	2.42	5.42	2.60	1.94	4.73	14.96	2.38	2.51	1.22
Brewery	3.79	3.94	3.81	3.58	3.15	1.10	3.62	13.47	3.23	3.08	1.69
Burning Peat	3.81	5.38	2.79	4.77	3.54	1.54	4.47	12.99	2.96	3.11	1.67
Burnt Wood	3.90	5.80	2.18	5.32	3.78	1.98	4.47	12.44	3.04	3.24	2.01
Cannon	2.78	6.40	1.72	5.90	2.74	2.60	5.10	15.66	2.22	2.54	1.34
Carbolic Soap	3.61	5.10	3.12	4.51	2.92	1.33	4.63	14.33	2.94	2.94	1.05

Casbah	3.86	4.70	3.86	3.88	3.44	1.26	4.45	13.93	3.16	3.02	1.63
Cedar	3.68	4.12	3.36	4.28	3.34	1.12	4.06	13.21	2.84	3.08	1.74
Church Incense	4.08	4.20	3.64	3.94	3.38	1.12	4.12	13.10	3.22	3.06	1.62
Cloisters	3.41	4.57	3.14	4.33	3.00	1.22	4.40	13.93	2.78	2.61	1.53
Clove Oil	3.04	5.02	2.58	4.60	2.61	1.70	4.43	15.10	2.47	2.46	1.24
Clover	4.13	4.02	4.54	3.14	3.41	1.02	4.00	13.28	3.26	3.16	1.63
Club	3.12	4.49	3.20	4.04	2.78	1.29	4.33	14.38	2.65	2.49	1.34
Coal Pit	3.49	6.49	1.67	5.96	2.33	2.33	5.08	16.29	2.14	2.49	1.26
Coal Soot	4.27	5.58	2.46	4.88	3.69	1.88	4.45	12.19	3.44	3.48	2.17
Cuban Cigar Smoke	3.37	4.52	3.47	4.04	2.59	1.35	4.60	15.13	2.61	2.61	1.05
Dentist	3.66	5.30	2.88	4.82	3.16	1.68	4.69	13.83	3.00	2.90	1.48
Earthy	4.76	4.88	3.80	3.76	3.76	1.20	4.55	11.68	3.52	3.22	1.78
Eau de Cologne	4.92	4.92	4.43	3.45	3.78	1.08	4.33	11.59	3.94	3.65	1.91
Eucalyptus	5.88	5.42	4.73	3.13	4.96	1.06	4.49	9.90	4.25	4.63	2.45
Fabric Softener	4.90	4.27	4.67	3.02	4.29	1.29	3.75	10.60	4.18	4.20	2.11
Farmyard	3.86	5.40	2.46	5.04	3.76	2.06	4.27	12.24	3.14	3.10	1.57

Firework	3.71	5.77	2.32	5.40	2.81	2.06	4.85	14.73	2.56	2.79	1.59
Football Pitch	4.52	5.00	3.90	3.75	3.81	1.19	4.19	12.17	3.67	3.38	1.85
Forest	4.14	4.04	4.18	3.46	3.62	1.06	3.84	12.55	3.40	3.34	1.71
Fox	4.30	4.62	3.76	3.84	3.50	1.08	4.04	12.32	3.18	2.94	1.75
Frosty	4.00	3.62	4.49	3.55	3.70	1.30	3.98	13.93	3.60	3.49	1.77
Gambia	4.58	4.13	4.21	3.13	3.60	1.08	3.79	11.34	3.38	3.25	1.83
Garden Shed	4.04	4.56	3.81	3.98	3.46	1.10	4.23	12.81	3.17	3.06	1.59
Ginseng	3.94	5.06	3.02	4.41	3.10	1.43	4.27	14.59	3.00	2.55	1.58
Grass/Hay	4.98	4.89	4.04	3.72	4.09	1.02	4.17	11.18	3.83	3.55	2.01
Havana Cigar	4.76	4.46	4.28	3.80	3.86	1.16	4.19	11.57	3.61	3.48	1.80
Hawthorn	4.37	4.14	4.39	3.45	3.35	1.00	4.02	11.12	3.61	3.37	1.94
Heather/Bracken	5.12	4.72	4.60	3.12	4.46	1.16	3.90	10.96	3.98	3.94	2.03
Honeysuckle	3.98	3.77	4.00	3.45	3.42	1.17	3.87	12.93	3.54	3.25	1.68
Hospital Modern Day	5.06	4.86	4.32	3.44	4.10	1.00	4.27	11.36	4.00	3.66	2.02
Hot Stuff Male	4.62	4.26	4.76	3.00	4.06	1.36	3.94	12.81	3.90	3.86	1.75
Hunter	3.73	3.85	4.13	3.52	3.42	0.92	3.81	13.79	3.48	3.27	1.45

Lavender	4.84	4.73	4.67	3.25	4.51	1.16	4.08	10.84	3.78	4.20	2.28
Leather	3.88	5.12	3.24	4.31	3.35	0.96	4.54	12.96	3.02	3.02	1.83
Leather Cream	4.37	3.92	3.55	3.59	3.96	1.10	4.10	11.56	3.27	3.41	1.50
Leather/Hide	3.60	4.28	3.55	4.00	3.13	1.09	4.35	14.14	2.85	2.89	1.46
Leaves	4.80	4.98	4.02	3.71	4.27	1.20	3.94	11.33	3.90	3.84	2.04
Lemon Cream	5.64	5.10	5.44	2.58	4.96	1.60	3.71	10.77	4.58	4.62	2.51
Mahogany	3.27	4.55	2.82	4.67	2.50	1.47	4.02	14.85	2.72	2.23	1.18
Man-o'-War	3.14	6.10	1.86	5.73	2.44	2.38	4.82	15.10	2.18	2.86	1.20
Garlic	5.10	6.51	2.06	5.51	4.47	2.18	4.21	10.91	3.94	4.00	2.30
Ginger	3.39	5.22	3.10	4.39	3.00	1.39	4.78	13.93	2.57	2.84	1.66
Gingerbread	3.69	4.04	3.51	3.79	2.79	1.09	4.19	13.74	2.73	2.60	1.59
Grapefruit	4.22	4.96	4.18	3.63	3.88	1.08	4.24	12.77	3.61	3.33	2.06
Hazelnut	4.40	5.40	3.25	4.31	3.35	1.79	4.27	11.51	3.23	3.19	1.95
Herring	4.02	4.64	2.88	4.46	3.76	1.68	3.90	11.50	3.22	3.32	1.99
Honey	4.50	4.10	4.24	3.34	3.70	1.16	3.80	10.22	3.38	3.36	2.07
Ice Cream	3.13	3.04	3.71	3.71	2.56	1.21	3.58	14.29	2.63	2.50	1.55

Iced Lemon	5.71	4.71	5.63	2.40	4.69	1.96	4.21	8.05	4.15	4.47	2.34
Irish Cream	4.63	4.31	4.02	3.58	3.92	1.10	4.25	14.62	3.50	3.56	2.10
Jelly Beans	3.75	3.73	4.40	3.11	2.85	1.23	3.81	12.07	2.96	2.77	1.74
Lemon, Eucalyptus & Mint	5.67	4.78	4.76	3.00	4.67	1.12	3.88	10.07	4.67	4.45	2.59
Lime	5.70	5.06	5.16	2.98	5.18	1.40	3.80	8.28	4.38	4.88	2.73
Liquorice	5.14	4.72	3.56	4.12	4.40	1.28	4.04	12.59	3.52	3.88	2.22
Lychee	4.38	4.50	4.44	3.33	3.75	1.56	3.90	11.77	3.25	3.19	1.75
Malted Barley	3.20	3.58	3.70	3.82	2.75	1.34	4.02	13.57	2.74	2.74	1.65
Mango & Sweet Orange	5.02	4.33	5.12	2.80	4.29	1.49	3.92	9.17	3.96	4.13	2.28
Mango Delight	4.40	4.42	3.90	3.53	3.70	1.70	3.92	12.60	3.20	3.62	1.83
Marzipan	6.12	5.27	4.96	3.00	5.14	1.65	4.22	7.88	4.00	4.90	2.73
Melon	5.37	4.55	5.39	2.92	3.96	1.67	4.35	11.50	3.71	3.90	2.20
Mixed Spice	4.42	4.88	4.04	3.92	3.72	1.40	4.45	12.17	3.50	3.42	1.97
Mulled Wine	5.08	3.94	5.14	2.68	3.94	1.46	3.84	11.12	4.26	3.74	2.01
Onion	4.66	6.58	2.00	6.02	4.22	2.28	4.46	11.34	3.82	4.20	2.28
Orange & Cinnamon	5.28	4.72	5.10	2.80	4.34	1.54	3.86	10.49	4.20	4.20	2.33

Orange (Seville)	5.47	4.45	5.20	2.61	4.69	1.61	3.86	9.47	4.33	4.47	2.30
Parma Violets	4.73	3.85	4.94	2.71	3.92	1.40	3.92	11.89	3.46	3.65	1.97
Passion Fruit	4.69	4.20	4.82	3.00	4.14	1.55	4.04	10.93	3.71	3.55	1.99
Peach Flesh	5.33	3.98	5.20	2.55	4.14	1.41	3.80	10.43	3.72	4.02	2.34
Peach Schnapps	5.73	4.49	5.65	2.59	4.57	1.69	3.65	8.77	4.06	4.22	2.40
Peanut	3.85	4.96	2.74	4.74	3.32	1.85	4.38	12.52	2.85	3.09	2.01
Pear	5.82	5.16	4.40	3.56	5.06	1.40	4.14	9.59	4.38	4.68	2.62
Pear Drops	5.66	5.00	4.74	3.08	4.56	1.58	4.12	8.89	4.02	4.16	2.28
Peppermint	5.56	5.18	4.64	3.16	4.98	1.08	3.92	10.66	4.51	4.70	2.55
Pineapple	4.69	4.35	4.53	3.24	3.96	1.31	4.10	10.43	3.65	3.69	2.36
Potato	4.47	5.41	2.82	4.78	3.84	1.80	4.45	12.94	3.69	3.59	2.11
Raspberry	3.88	3.90	4.12	3.40	2.86	1.32	4.18	13.28	2.86	2.86	1.80
Rhubarb	4.22	4.80	3.92	3.98	3.08	1.56	4.18	12.60	3.00	2.92	1.72
Rosemary	4.53	4.96	3.84	3.82	3.94	0.98	4.27	11.51	3.31	3.53	2.32
Rum	4.30	5.64	2.90	4.94	3.50	1.54	4.52	14.28	2.98	3.28	1.70
Sage	4.96	5.24	4.12	3.58	4.24	0.92	4.38	10.00	3.46	3.69	2.28

Shea & Butter	5.43	4.76	4.90	2.96	4.69	1.22	3.84	9.45	4.69	4.45	2.16
Spearmint	5.90	4.96	5.08	2.60	5.72	1.48	3.56	7.83	5.57	5.40	2.71
Strawberry	5.37	4.24	5.37	2.45	4.69	1.82	3.76	10.64	4.06	4.41	2.22
Sweet Sherry	4.58	5.46	3.16	4.62	3.88	1.48	4.54	12.63	2.98	3.50	2.03
Tea Leaf	4.02	4.59	3.80	3.96	3.10	1.06	4.24	13.27	2.98	2.98	1.72
Toffee Apple	4.10	4.60	4.04	3.78	2.96	1.12	4.22	13.24	3.33	3.10	1.63
Vanilla	3.66	3.60	3.76	3.60	3.00	1.12	3.96	13.52	2.88	2.65	1.81
Watermelon	5.34	4.54	5.52	2.50	4.48	1.80	3.76	9.50	4.08	4.36	2.24
Whisky	3.61	5.02	2.59	4.71	3.08	1.65	4.54	13.83	2.84	2.86	1.76
Tomato Plant	4.88	4.78	4.27	3.37	4.02	0.84	3.98	11.45	4.04	3.65	2.01
Menthol	3.30	3.30	3.68	3.74	2.78	1.00	3.78	14.54	2.72	2.70	1.34
Methane	3.57	5.71	1.71	5.65	3.22	2.29	4.49	13.80	2.55	3.13	1.72
Mountain Heather	4.80	4.12	4.47	3.00	4.22	1.04	3.71	10.64	3.73	3.76	2.45
Mouse	3.36	5.06	2.70	4.70	2.78	1.50	4.76	14.50	2.62	2.38	1.53
Myrrh	3.71	4.31	3.86	3.88	3.22	0.80	4.12	12.85	3.08	2.92	1.63
Nag Champa	3.10	3.41	3.73	3.78	2.35	0.77	4.08	16.08	2.57	2.29	1.28

New Car	4.18	3.43	4.61	3.24	3.61	0.86	3.65	12.00	3.53	3.20	2.01
Nutmeg	3.14	4.32	3.24	4.34	2.60	1.32	4.60	15.22	2.64	2.50	1.25
Oak	3.37	4.51	2.90	4.47	2.78	1.43	4.33	13.61	2.90	2.65	1.76
Old Drifter (Ship)	3.39	5.27	2.31	5.10	3.06	1.82	4.49	14.38	2.67	2.90	1.99
Old Inn	3.18	4.12	2.86	4.32	2.50	1.50	4.18	14.41	2.34	2.42	1.46
Old Smithy	4.60	4.72	3.98	3.65	3.94	1.06	4.24	11.02	3.54	3.28	1.99
Out At Sea	4.65	4.67	3.67	3.90	3.33	1.08	4.10	11.07	3.50	3.29	1.99
Ozone	3.92	4.38	3.68	3.82	3.46	1.28	4.35	11.35	3.34	3.12	1.75
Patchouli	3.55	5.06	3.02	4.43	2.67	1.31	4.78	13.89	2.85	2.47	1.62
Peat	3.31	3.27	3.94	3.63	2.78	0.76	3.76	14.76	2.73	2.38	1.63
Pencils	3.02	3.86	3.33	3.94	2.73	1.04	3.98	14.78	2.61	2.57	1.45
Pine/Heather/Peat	4.42	4.64	3.92	3.84	3.50	1.08	4.56	12.36	3.30	3.16	1.83
Polish/Wax	3.94	4.24	4.20	3.51	3.35	1.18	4.18	12.55	3.16	3.20	1.89
Practical Man	4.86	4.48	4.74	3.16	4.10	1.26	4.28	12.56	3.96	3.72	1.98
Racing Car	3.38	6.32	1.88	5.42	2.90	2.40	4.62	14.18	2.53	2.66	1.45
Rockpools	3.61	3.35	4.04	3.65	2.78	1.14	3.55	13.86	2.69	2.84	1.47

Roselle	4.06	4.00	4.20	3.24	3.24	1.20	4.06	12.23	3.12	3.00	1.88
Rum Barrel	3.10	5.18	2.68	4.66	2.54	1.56	4.66	15.65	2.46	2.50	1.26
Sandalwood	2.73	2.96	3.73	3.69	2.27	0.69	3.83	14.90	2.50	2.13	1.17
Sea Breeze	4.46	4.38	4.14	3.62	3.92	1.12	4.04	11.29	3.94	3.76	2.18
Sea Mineral	5.20	4.39	4.53	3.33	4.41	1.18	4.06	10.90	4.06	3.79	2.23
Sea Shore	2.96	5.20	2.20	5.10	2.29	1.84	4.81	15.86	2.27	2.08	1.53
Smugglers	4.56	5.52	2.12	5.50	3.78	2.36	3.92	12.89	3.34	3.70	2.17
Sports Changing Room	4.16	5.12	3.43	4.06	3.49	1.06	4.69	12.12	3.10	3.00	1.84
Sports Rub	5.60	5.52	4.21	3.29	5.19	1.08	4.19	8.52	4.45	4.38	2.69
Stars Dressing Room	4.38	4.06	4.49	3.30	3.89	1.17	4.04	11.33	3.53	3.43	1.95
Sun tan lotion	5.16	4.22	5.00	2.80	4.64	1.52	4.10	8.92	4.42	3.90	2.22
Tarmac	3.10	6.19	1.96	5.67	2.64	2.29	5.21	14.86	2.38	2.85	1.86
Tea Tree Oil	5.20	5.14	3.67	3.80	4.04	1.18	4.47	10.72	3.65	3.86	2.20
Tobacco Leaf	3.75	4.73	3.10	4.44	3.00	1.35	4.33	13.74	3.13	2.77	1.88
Toothpaste	3.80	4.94	3.37	4.22	3.20	1.45	4.27	12.76	3.27	3.12	1.84
Train Smoke	3.96	6.06	2.18	5.30	3.37	2.10	4.68	12.31	2.92	2.94	2.39

Trophy Room	2.94	4.46	2.98	4.60	2.17	1.35	4.58	15.18	2.25	2.15	1.25
Turpentine	4.33	4.59	3.65	3.69	3.45	0.96	4.24	12.41	3.39	3.13	1.78
Tyres	4.00	5.80	2.18	5.10	3.71	1.86	5.06	12.61	2.86	3.08	1.99
Victorian Street	2.98	5.59	1.86	5.49	2.39	2.31	4.57	14.92	2.31	2.27	1.62
Washday	4.78	4.86	3.96	3.61	4.02	1.08	3.94	10.67	3.86	3.57	1.95
Washing Up Liquid	5.65	4.71	5.33	2.65	4.80	1.75	3.78	9.64	4.46	4.35	2.45
Wood Chip	2.94	3.96	3.80	3.98	2.42	1.16	4.12	14.60	2.66	2.28	1.18
Woodsmoke	4.28	5.00	2.94	4.88	3.62	1.50	4.28	14.54	3.06	3.22	2.11
Ylang, Jasmine & Myrrh	4.80	4.56	4.60	3.18	4.02	0.96	4.22	10.33	4.30	3.74	2.24
Aftershave	4.96	4.24	4.74	3.02	4.34	1.06	4.04	12.19	4.22	3.78	2.31
Soap Suds	4.90	4.14	4.37	3.22	4.63	1.10	3.53	9.60	4.29	4.10	2.54
Rubber	4.12	5.80	2.22	5.39	3.90	1.86	4.98	12.06	3.17	3.16	2.15

8.2 Appendix B: Table of normative data for odours used in Chapter 3.

Odour	Group	Experiment 1 & 2	Experiment 3	Experiment 4	Verbalisability	Familiarity	Intensity	Pleasantness	Hedonic Strength
Cinder Toffee	Low	X		X_1	1.02	3.02	3.37	3.59	1.04
Carbolic Soap	Low	X	X	X_2	1.05	3.61	5.10	3.12	1.33
Cuban Cigar Smoke	Low	X	X	X_1	1.05	3.37	4.52	3.47	1.35
Sandalwood	Low	X		X_1	1.17	2.73	2.96	3.73	0.69
Wood Chip	Low	X		X_2	1.18	2.94	3.96	3.80	1.16
Nutmeg	Low	X		X_2	1.25	3.14	4.32	3.24	1.32
Trophy Room	Low			X_1	1.25	2.94	4.46	2.98	1.35
Rum Barrel	Low		X	X_2	1.26	3.10	5.18	2.68	1.56
Nag Champa	Low	X*		X_2	1.28	3.10	3.41	3.73	0.77
Old Inn	Low			X_1	1.46	3.18	4.12	2.86	1.50
Rockpools	Low			X_2	1.47	3.61	3.35	4.04	1.14
Mouse	Low		X		1.53	3.36	5.06	2.70	1.50

Sea Shore	Low		X		1.53	2.96	5.20	2.20	1.84
Patchouli	Low		X		1.62	3.55	5.06	3.02	1.31
Ginger	Low			X_1	1.66	3.39	5.22	3.1	1.39
Blackcurrant	High			X	2.44	5.67	4.85	5.48	1.73
Eucalyptus	High	X	X		2.45	5.88	5.42	4.73	1.06
Lemon Cream	High			X	2.51	5.64	5.10	5.44	1.60
Aniseed Balls	High			X	2.61	5.88	5.40	3.98	1.50
Garden Mint	High	X	X		2.62	5.84	5.08	5.30	1.66
Pear	High	X*	X	X	2.62	5.82	5.16	4.40	1.40
Sports Rub	High	X	X		2.69	5.60	5.52	4.21	1.08
Spearmint	High	X		X	2.71	5.90	4.96	5.08	1.48
Lime	High	X	X	X	2.73	5.70	5.06	5.16	1.40
Marzipan	High	X	X		2.73	6.12	5.27	4.96	1.65

* Denotes the odour used as a buffer item in the first two trials of an n-back testing sequence.

8.3 Appendix C: The Self-Assessment Manikin (SAM) Scale Instructions

This scale requires you to rate how each odour makes you feel whilst you are smelling it. There are no right or wrong answers, so simply respond as honestly as you can.

The SAM scale has 3 sets of 5 pictures, with each set arranged in a continuum. You will use these figures to rate how each odour made you feel. The three kinds of feelings that the SAM shows is Happy vs. Unhappy, Excited vs. Calm, and Controlled vs. In-control.

The **left** of the happy vs unhappy scale should be clicked if when smelling the odour you felt completely 'happy, pleased, satisfied, contented, or hopeful'. The **right** of the scale should be clicked if you felt completely 'unhappy, annoyed, unsatisfied, melancholic, despaired, bored'.

The **left** side of the excited vs calm scale should be clicked if you felt stimulated, excited, frenzied, jittery, wide-awake, aroused. The **right** side should be clicked if you felt completely 'relaxed, calm, sluggish, dull, sleepy, and unaroused'.

The final scale is the feeling of being controlled or being in-control. The **left** side of the scale should be clicked if you have feelings characterised as completely 'controlled, influenced, cared-for, awed, submissive, and guided'. The **right** side of scale reflects feeling completely 'controlling, influential, in control, important, dominant, and autonomous'.

Some of the odours may prompt emotional experiences; others may seem relatively neutral. Your rating of each odour should reflect your immediate personal experience, and no more. Please rate each one as you actually felt as you smelt it.

8.4 Appendix D: Additional analyses for Chapter 4

Hit rate analysis. Analysis of hit rates revealed a similar pattern of results observed in the *A'* score analysis. That is, there was evidence for a main effect of task type, $F(3, 165) = 14.89$, $p < .001$, $\eta_p^2 = .21$ (model vs the null, $BF_{10} > 1,000$), and anecdotal evidence for better target responding for verbalisable odours than the low verbalisability odours ($p = .038$, $BF_{10} = 2.26$). There was, however, evidence for no difference between the visual 2-back task and verbalisable odours ($p = .258$, $BF_{10} = 0.27$), nor with low verbalisability odours ($p = .333$, $BF_{10} = 0.23$). Between the 2-back task hit rates and the verbal 3-back task there was evidence for difference across all comparisons. That is, hits were higher than the verbal task for the visual 2-back ($p < .001$, $BF_{10} > 1,000$), the verbalisable odour 2-back ($p < .001$, $BF_{10} > 1,000$), and, in contrast to the *A'* findings, also in the low verbalisability odour 2-back ($p < .001$, $BF_{10} = 106.54$).

Close-lure false alarm rate analysis. False alarms for only *close-lure* trials (n+1 and n-1) were analysed as an index of working memory that is not influenced by familiarity-based recognition (Harbison et al., 2011). The data supported a main effect of task, $F(3, 165) = 11.52$, $p < .001$, $\eta_p^2 = .17$ (model vs the null, $BF_{10} > 1,000$). Paired comparisons revealed support for greater close-lure false alarms in low verbalisability odour memory than verbalisable odour memory ($p = .005$, $BF_{10} = 21.17$). There was also evidence for a difference between low verbalisability odour false alarms and visual memory false alarms ($p < .001$, $BF_{10} > 1,000$), with greater false alarms in the odour task. However, there was only anecdotal support for a difference between low verbalisability odour memory and verbal memory ($p = .079$, $BF_{10} = 1.83$). Furthermore, whilst there was evidence for a difference between verbal and visual memory ($p = .007$, $BF_{10} = 527.91$), with higher false alarms in the verbal task, there was evidence against a difference between verbalisable odour false alarms and verbal false alarms ($p = .359$, $BF_{10} = 0.22$), and only anecdotal evidence between verbalisable odours and visual task false alarms (p = .018, BF_{10} = 2.18).

8.5 Appendix E. Odour normative data for stimuli used in Chapter 5.

Odour	Experiment	Fam.	Int.	Pleas.	Irr.	Verb.
Boiler Room	1	3.04	5.86	2.42	5.42	1.22
Man-o'-War	1	3.14	6.10	1.86	5.73	1.20
Racing Car	1	3.38	6.32	1.88	5.42	1.45
Ginger	1	3.39	5.22	3.10	4.39	1.66
Coal Pit	1	3.49	6.49	1.67	5.96	1.26
Blue Cheese	1	3.51	5.66	1.89	5.60	1.66
Farmyard	1	3.86	5.40	2.46	5.04	1.57
Casbah	1	3.86	4.70	3.86	3.88	1.63
Coriander	1	3.87	4.76	3.98	3.87	1.51
Frosty	1	4.00	3.62	4.49	3.55	1.77
Tea Leaf	1	4.02	4.59	3.80	3.96	1.72
Basil	1	4.06	4.43	4.12	3.37	1.64
Toffee Apple	1	4.10	4.60	4.04	3.78	1.63
Clover	1	4.13	4.02	4.54	3.14	1.63
Rhubarb	1	4.22	4.80	3.92	3.98	1.72
Lychee	1	4.38	4.50	4.44	3.33	1.75
Pine/Heather/Peat	1	4.42	4.64	3.92	3.84	1.83
Cinnamon	1	4.54	5.00	4.44	3.52	1.79
Hot Stuff Male	1	4.62	4.26	4.76	3.00	1.75
Black Pepper	1	4.88	5.04	4.24	3.66	1.77
Aniseed Balls	2	5.88	5.40	3.98	3.54	2.61
Pear	2	5.82	5.16	4.40	3.56	2.62
Pear Drops	2	5.66	5.00	4.74	3.08	2.28
Fruit Punch	2	5.45	4.18	5.29	2.49	2.45
Peach Flesh	2	5.33	3.98	5.20	2.55	2.34

Tea Tree Oil	2	5.20	5.14	3.67	3.80	2.20
Liquorice	2	5.14	4.72	3.56	4.12	2.22
Garlic	2	5.10	6.51	2.06	5.51	2.30
Mango & Sweet Orange	2	5.02	4.33	5.12	2.80	2.28
Sage	2	4.96	5.24	4.12	3.58	2.28
Chocolate (Mint)	2	4.90	5.48	4.12	3.84	2.24
Soap Suds	2	4.90	4.14	4.37	3.22	2.54
Fish	2	4.74	5.78	1.66	5.66	2.57
Pineapple	2	4.69	4.35	4.53	3.24	2.36
Onion	2	4.66	6.58	2.00	6.02	2.28
Irish Cream	2	4.63	4.31	4.02	3.58	2.10
Rosemary	2	4.53	4.96	3.84	3.82	2.32
Coal Soot	2	4.27	5.58	2.46	4.88	2.17
Rubber	2	4.12	5.80	2.22	5.39	2.15
Train Smoke	2	3.96	6.06	2.18	5.30	2.39

Fam = Familiarity, Int = Intensity, Pleas = Pleasantness, Irr = Irritability, Verb = Verbalisability