Journal of Experimental Psychology: General
2017, Vol. 146, No. 9, 1266–1285

0096-3445/17/$12.00 http://dx.doi.org/10.1037/xge0000333

Shifting Between Mental Sets: An Individual Differences Approach to Commonalities and Differences of Task Switching Components

Claudia C. von Bastian
Bournemouth University and University of Zurich

Michel D. Druey
University of Zurich and University of Gloucestershire

Switching between mental sets has been extensively investigated in both experimental and individual differences research using a wide range of task-switch paradigms. However, it is yet unclear whether these different tasks measure a unitary shifting ability or reflect different facets thereof. In this study, 20 task pairs were administered to 119 young adults to assess 5 proposed components of mental set shifting: switching between judgments, stimulus dimensions, stimulus–response mappings, response sets, and stimulus sets. Modeling latent factors for each of the components revealed that a model with 5 separate yet mostly correlated factors fit the data best. In this model, the components most strongly related to the other latent factors were stimulus–response mapping shifting and, to a lesser degree, response set shifting. In addition, both factors were statistically indistinguishable from a second-order general shifting factor. In contrast, shifting between judgments as well as stimulus dimensions consistently required separate factors and could, hence, not fully be accounted for by the general shifting factor. Finally, shifting between stimulus sets was unrelated to any other shifting component but mapping shifting. We conclude that tasks assessing shifting between mappings are most adequate to assess general shifting ability. In contrast, shifting between stimulus sets (e.g., as in the Trail Making Test) probably reflects shifts in visual attention rather than executive shifting ability.

Keywords: task switching, individual differences, executive functions, unity and diversity

Supplemental materials: http://dx.doi.org/10.1037/xge0000333.supp

It is common experience that we rarely focus on only one activity at the time uninterruptedly. Rather, most of our daily life involves several things going on in quick succession or even concurrently. For example, while driving a car, we may simultaneously listen to music, quickly shift our attention from the street to the navigation system, and back on the road in an attempt to orient in an unfamiliar city. The mental processes behind such task switching and multitasking are obviously complex and have been one of the main topics in cognitive research of the last decades. Terms as, for instance, *executive functions*, *cognitive control*, or *mental flexibility* have been used to refer to the set of "shifting abilities" required in such situations, and many different paradigms have been developed to experimentally investigate these skills (e.g., Logan, 1985; Pashler, 2000).

Set shifting abilities have also been of central interest in individual differences studies and studies on so-called frontal-lobe functions (e.g., Alvarez & Emory, 2006; Miyake et al., 2000). In these contexts, tasks have frequently been used that already have a long tradition in the assessment of frontal-executive functioning such as the Wisconsin Card Sorting Test (WCST) and its derivatives, or the Trail Making Test (TMT).

In the present study, we try to bridge the two fields by using an individual differences approach and tasks designed in the field of experimental task-switching research to disentangle the commonalities and differences in switching our attention between different representational sets in an attempt to search for a common underlying shifting ability. Before outlining the details of our study, though, we first give a short overview of the literature regarding the typical tasks used in experimental studies and the question of what characterizes task (sets). Next, we consider individual differences studies of mental set shifting. Here, we primarily focus on the tasks that have been used to assess shifting abilities. Integrating these two parts, we then sketch a first heuristic model which guided the design of the present study.

Task Switching and the Conceptualization of Task Sets

Since the seminal studies by Allport, Styles, and Hsieh (1994), and Rogers and Monsell (1995), several task-switch

This article was published Online First June 8, 2017.

Claudia C. von Bastian, Cognition and Cognitive Neuroscience Research Centre, Department of Psychology, Bournemouth University, and Department of Psychology, University of Zurich; Michel D. Druey, Department of Psychology, University of Zurich, and School of Natural and Social Sciences, Psychology, University of Gloucestershire.

Michel D. Druey is now at the Department of Psychology, University of Konstanz.

We thank Sabrina Guye, Jessica Kohlberg, and Jessica Oschwald for their assistance with collecting the data. We also thank Klaus Oberauer for his comments on a draft of this article.

The ideas and data presented in this article have not been presented elsewhere. Both authors contributed equally to this article.

Correspondence concerning this article should be addressed to Claudia C. von Bastian, Department of Psychology, Bournemouth University, Talbot Campus, Fern Barrow, Poole BH12 5BB, United Kingdom. E-mail: cvonbastian@bournemouth.ac.uk

paradigms have been established. The most extensively used is the task-cuing paradigm, in which participants perform a random sequence of at least two tasks. The relevant task on a given individual trial is indicated by a cue that is presented some time before the onset of the actual target stimulus (cue-target interval [CTI]). Usually, the stimuli presented afford both tasks (i.e., are bivalent) as, for instance, in the case where digits are presented that have to be classified according to either their parity or magnitude. Also, in most studies, the same responses are used for both tasks; hence there is a response-set overlap between tasks (e.g., even digits and digits smaller than 5 require pressing a left key, and odd digits and digits larger than 5 require pressing a right key). Figure 1 depicts an exemplar two-trial switch sequence in this type of paradigm.

Typically, participants respond faster to target stimuli in trials in which the previous task repeats than to target stimuli in trials in which the task switches (e.g., Koch, 2001; Meiran, 1996; Rogers & Monsell, 1995). The difference in response time (RT) or error rate between repetition and switch trials in mixed-task blocks are the switch costs. Despite being the main focus of task-switch research, there is still much debate about the processes contributing to the switch costs (for reviews, see Kiesel et al., 2010; Vandierendonck, Liefooghe, & Verbruggen, 2010). Most researchers agree though on these costs being caused by multiple processes. Attempts to model the temporal dynamics in task switch situations (e.g., Altmann & Gray, 2008; Logan & Bundesen, 2003; Mayr & Kliegl, 2000; Meiran, 2000a; Meiran, Kessler, & Adi-Japha, 2008; Rubinstein, Meyer, & Evans, 2001; Rushworth, Passingham, & Nobre, 2002, 2005) typically emphasize two main sources of switch costs: (the lack of) advance task preparation and carry-over effects from the previous task (i.e., proactive interference).

The conceptualizations of what constitutes a task (set) are similarly diverse. Mayr and Keele (2000), for example, suggested that "task sets are assumed to specify the configuration of perceptual, attentional, mnemonic, and motor processes critical for a particular task goal" (p. 5), thereby highlighting the relevance of cognitive parameter settings in situations with ambiguous stimulus input, but where a specific goal is to be reached. Such processing-related conceptualizations of task sets are typical for formal models of task switching (e.g., in the parallel distributed processing-model [PDP-model] by Gilbert & Shallice, 2002; in executive control of the theory visual attention [i.e., ECTVA], Logan & Gordon, 2001; in control by action representation and input selection [i.e., CARIS], Meiran et al., 2008).

Other attempts to specify the concept of task sets focus on structural rather than functional aspects. Kleinsorge and Heuer (1999; Kleinsorge, 2004), for example, suggested a hierarchical structure of task representations with four main task features. We consider this model in more detail here, as it is central to our study. The four main task features Kleinsorge and Heuer (1999; Kleinsorge, 2004) suggested are judgments, dimensions, mappings, and responses. On the top level, the task set hierarchy encompasses the judgments (i.e., parity and magnitude in the above example). Subordinate to the judgments are the stimulus dimensions (i.e., the digit values) and the stimulus–response mappings (i.e., even digits and digits smaller than 5 are assigned to pressing a left key, and odd digits and digits larger than 5 are assigned to pressing a right key). At the lowest level of the task hierarchy are the responses (i.e., the key presses and their mental representations).

In the present study, we adopted the structural task-set conceptualization provided by Kleinsorge and Heuer (1999; Kleinsorge, 2004). Different than experimental research on the temporal structure of basic selection processes contributing to the switch costs (e.g., see Meiran, 2000a; Meiran et al., 2008; Rushworth et al., 2002, 2005), we were interested in how different shifting components are related from an individual differences perspective, thereby identifying which elements of a task possibly require the contribution of a general shifting ability. For this purpose, we classified tasks that have been used in previous experimental task-switching research into categories according to the task feature that had to be switched. Figure 2 illustrates the involved representational changes in task-switch trials for each of these task

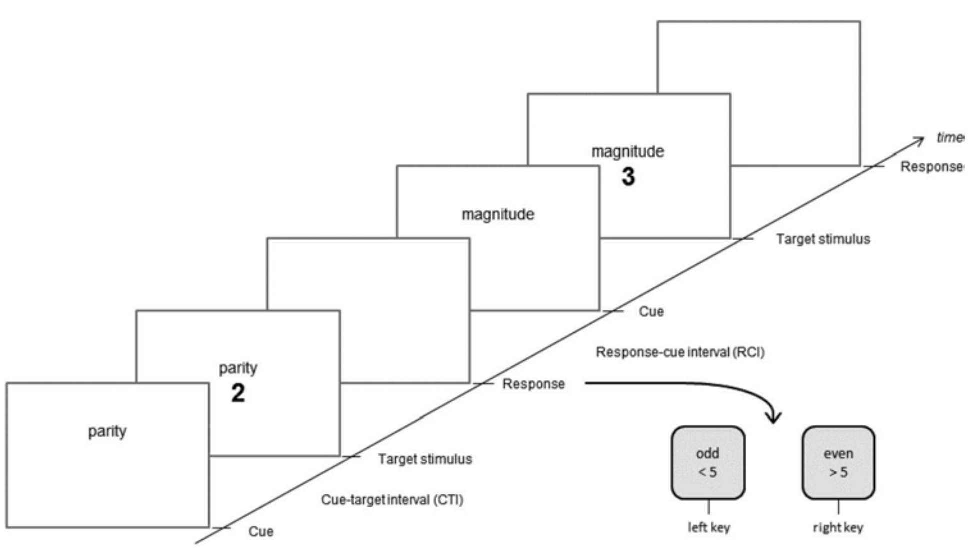
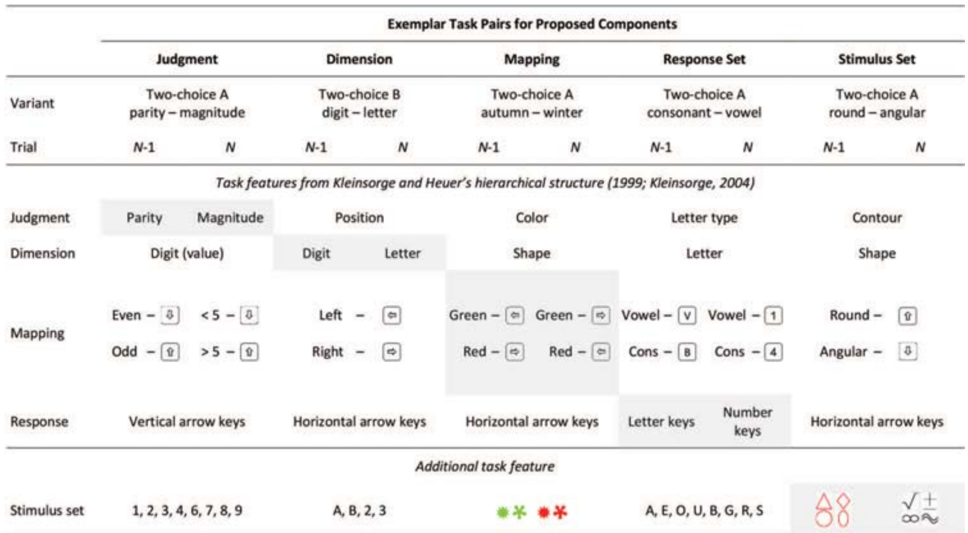


Figure 1. Exemplar two-trial task-switch sequence in a typical cued task-switch experiment.



	Judgment		Dimension		Mapping		Response Set		Stimulus Set	
	Exemplar Task Pairs for Proposed Components									
Variant	Two-choice A parity – magnitude		Two-choice B digit – letter		Two-choice A autumn – winter		Two-choice A consonant – vowel		Two-choice A round – angular	
Trial	*N*-1	*N*	*N*-1	*N*	*N*-1	*N*	*N*-1	*N*	*N*-1	*N*
	Task features from Kleinsorge and Heuer's hierarchical structure (1999; Kleinsorge, 2004)									
Judgment	Parity	Magnitude	Position		Color		Letter type		Contour	
Dimension	Digit (value)		Digit	Letter	Shape		Letter		Shape	
Mapping	Even – ⇩ Odd – ⇧	< 5 – ⇩ > 5 – ⇧	Left – ⇦ Right – ⇨		Green – ⇦ Red – ⇨	Green – ⇨ Red – ⇦	Vowel – V Cons – B	Vowel – 1 Cons – 4	Round – ⇧ Angular – ⇩	
Response	Vertical arrow keys		Horizontal arrow keys		Horizontal arrow keys		Letter keys	Number keys	Horizontal arrow keys	
	Additional task feature									
Stimulus set	1, 2, 3, 4, 6, 7, 8, 9		A, B, 2, 3				A, E, O, U, B, G, R, S			

Figure 2. Example two-choice task pairs for the five proposed shifting components: the four task-set features of Kleinsorge and Heuer (1999; Kleinsorge, 2004), and the additional stimulus-set task feature. Depicted are the switch-transitions from one trial to the next (i.e., N-1 to N) on each task-set feature for the respective task pairs, with the critical switches for the respective component highlighted in gray. Note that, for methodological reasons, the exact stimulus–response mapping changes in all components except for dimension shifting. For all components, the two-choice task pair with the highest reliability is depicted (see text for more details). See the online article for the color version of this figure.

features as well as for an additional feature we included (see the following text).

Task Switching and a General Shifting Ability: Individual Differences Approaches

To our best knowledge, it has not yet been systematically investigated whether and, if so, which different types of shifting components of a task contribute to a common underlying shifting ability in an individual differences framework. However, a variety of shifting tasks has been used to assess the relationship between mental set shifting ability and other cognitive abilities such as inhibition, working memory, and fluid intelligence (e.g., Miyake & Friedman, 2012; Miyake et al., 2000; Oberauer, Süß, Schulze, Wilhelm, & Wittmann, 2000; Süß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002). In these studies, shifting ability has been characterized as switching "back and forth between multiple tasks, operations, or mental sets" (Miyake et al., 2000, p. 55). Miyake and colleagues operationalized shifting ability with three different task pairs which they considered different enough to not engage exactly the same operations to be switched, but similar enough in that all require shifting between mental sets. The tasks they used were the plus–minus task from Jersild (1927), the number–letter task from Rogers and Monsell (1995), and the global-local task (cf. R. Hübner, Futterer, & Steinhauser, 2001). The results showed that these tasks indeed loaded well on one shifting factor, and that this factor was clearly separable from two other executive functions ("Inhibition" and "Updating"). Moreover, the shifting factor predicted performance in a frequently used "frontal-lobe" task, the WCST.

Although these tasks share some crucial features as indicated by their correlations and the loading on one factor, they also differ in several respects when viewed from the perspective of the structural task-set framework proposed by Kleinsorge and Heuer (1999; Kleinsorge, 2004). For example, the judgment for the global and local shapes in the global-local task of Miyake et al. (2000) remained the same (i.e., naming the number of lines a shape is composed of) and shifting was only required for the level (global vs. local, i.e., the stimulus dimension). In contrast, in the number–letter task, participants switched between a letter judgment (vowel or consonant) and a digit judgment (odd or even) in a two-symbol display with one letter and one digit side-by-side.

In a similar vein, Oberauer and colleagues (2000; Süß et al., 2002) operationalized shifting with three different tasks with numerical, figural, and verbal material, respectively. The numerical task (previously used by Allport et al., 1994) required participants to switch between judging the value of a set of (identical) digits on a screen, and the number of digits presented. In the figural and the verbal tasks, participants had to decide on which side a round or angular shape was presented in a horizontally arranged two-stimulus display (figural task), or whether the left or right word in a two-word display belonged to a prespecified category. As the results showed, these three tasks were highly correlated and loaded on a common factor in a structural equation model, despite them covering different task features according to Kleinsorge and Heuer's (1999; Kleinsorge, 2004) task set framework. More specifically, whereas the numerical task required participants to shift between judgments, the judgments remained constant in the figural and verbal tasks. In these tasks, shifting was instead required between stimulus dimensions.

To capture these commonalities and differences between the tasks, the present study attempts to extend the Kleinsorge (2004) and Kleinsorge and Heuer (1999) framework. In the following, we outline how switching tasks used in the context of experimental and individual differences studies map on such an extension of Kleinsorge's framework in more detail.

Types of Tasks and a First Model

In experimental research, switching has been operationalized at all four representational levels (i.e., judgment, dimension, mapping, and response sets) of Kleinsorge and Heuer's (1999; Kleinsorge, 2004) framework. More specifically, experimental paradigms measured switching at the judgment level (as in the vast majority of task-switch studies; e.g., Altmann, 2004; Arrington & Logan, 2004; Dreisbach, Haider, & Kluwe, 2002; M. Hübner, Kluwe, Luna-Rodriguez, & Peters, 2004; Koch, 2001; Lien, Schweickert, & Proctor, 2003; Logan & Bundesen, 2003; Logan, Schneider, & Bundesen, 2007; Mayr & Kliegl, 2000; Meiran, 1996; Nieuwenhuis & Monsell, 2002; Oriet & Jolcœur, 2003; Rogers & Monsell, 1995; Schneider & Logan, 2006; Waszak, Hommel, & Allport, 2003), at the stimulus dimension level (e.g., R. Hübner et al., 2001; Mayr & Keele, 2000), at the level of the responses (in terms of response sets; e.g., Philipp & Koch, 2005), and at the level of (pure) stimulus–response mappings (e.g., Rushworth et al., 2002, 2005; Shaffer, 1965).

Similarly, the shifting tasks that have been used in individual differences research vary across studies (see Table 1 for a nonexhaustive overview), with the majority of tasks measuring judgment and dimension shifting. Interestingly, in many studies shifting has been operationalized by using the TMT (see Table 1). In the TMT's critical Part B, participants have to alternately connect the numbers 1 to 12 and the letters A to L by a line (i.e., 1A2B3C . . .). Such shifting between mere stimulus sets is not reflected in Kleinsorge and Heuer's (1999; Kleinsorge, 2004) task set framework. Therefore, we added stimulus set shifting as a fifth representational level to our model (see Figure 2).[1]

Because of methodological constraints, different shifting components are inevitably often concurrently present when switching from one task to another. For example, exact stimulus–response (SR) mapping shifts are immanently present in tasks measuring all other components, except only those requiring pure stimulus dimension shifts (Mayr, 2002; Mayr & Keele, 2000; see also Figure 2). Similarly, stimuli are necessarily bivalent (i.e., afford both tasks) in any shifting component other than the one measuring pure stimulus set shifting. Thus, measuring pure mapping and pure stimulus set shifting and including them as latent factors in the model allowed for examining how much influence these components exert on the (necessarily) less pure judgment, dimension, and response set shifting components. Furthermore, regarding the mapping shifting component, inclusion is also indicated for theoretical reasons. In his working memory (WM) model, Oberauer (2009) assumed that the costs observed under task-switch conditions are mainly due to the trial-by-trial (un-)loading of the stimulus–response mappings. Hence, according to this model, the mappings are a key component regarding mental set shifting ability.

Present Study

In this study, we used different task pairs from previous experimental task-switch research in an individual differences approach to answer the following research questions: (a) to what extent can the proposed five shifting components be constituted as separable factors and how are they interrelated, and (b) can these interrelations be accounted for by an underlying general shifting ability? For this purpose, for each of the five components, we adapted four tasks from the literature or, in cases where we did not find adequate tasks, created new ones.

According to previous evidence, we expected switch costs for all five components, possibly with the exception of stimulus set shifting, as this component presumably only requires shifting of visual-spatial and not of executive attention (cf. Miyake et al., 2000). Figure 3 illustrates the basic versions of a respective series of models we tested to answer our research questions. The measurement model (see Figure 3A) served to evaluate the separability of latent factors reflecting the five theoretically assumed shifting components and their correlations. The hierarchical models (see Figure 3B) served to test the notion of a general shifting ability accounting for variance across shifting components. For this purpose, we modeled a second-order general shifting factor with loadings from the latent component factors. Finally, the bifactor (sometimes also referred to as nested-factor) models (see Figure 3C) served to explore whether any of the shifting components constitutes variance over and above a general shifting factor. Therefore, here, we modeled component-specific factors in addition to a general shifting factor, allowing for the measures to load on both.

In addition, for each component, two tasks were selected that involved two response options (most common in experimental task switch research), one task with three response options, and one task with four response options (as there are also task-switch studies with more than two response options). Moreover, across the five components, we used both figural and verbal-numerical materials. This allowed us to test two alternative task-driven models based on the number of response options (2, 3, or 4) and the stimulus materials (verbal–numerical or figural).

Method

Participants

We aimed at collecting data from at least 120 participants. We determined the sample size based on previous individual differences studies by members of our laboratory (Oberauer et al., 2000; Oberauer, Süß, Wilhelm, & Wittmann, 2003; von Bastian & Oberauer, 2013; von Bastian, Souza, & Gade, 2016; see Discussion for a further consideration of power). We recruited 127 students from the University of Zurich, Switzerland, who either received course credit or 50 CHF in exchange for their participation. As our primary dependent variable was based on RTs to correct responses, we excluded eight participants who performed below the prede-

[1] Note that we use the term *stimulus set* only descriptively in the sense that participants shift between two different sets of stimuli. It is thus not related to the "S(timulus)-Set" concept used in Meiran's (2000a) framework, where S-Sets play a functional role (in terms of selective attention) for the biasing of response selection. In later work (Meiran et al., 2008), Meiran abandoned this terminology, using "input set" instead. The same restriction pertains to the "R(esponse)-Set" concept in Meiran (2000a), which was later changed to "action set" (Meiran et al., 2008). In contrast, we use the term *response set* in a structural-descriptive sense, referring to separate sets (chunks) of responses.

Table 1
Individual Differences Studies of Shifting (Nonexhaustive)

Study	Task	Shifting type
Adrover-Roig et al. (2012)	Brixton Test[a]	Judgment
	Madrid Card Sorting Test[a]	Judgment
	TMT	Stimulus set
Agostino et al. (2010)	Contingency naming task	Judgment
	TMT	Stimulus set
de Frias et al. (2009)	Brixton Test[a]	Judgment
	Color Trails Test[b]	Stimulus set
Fisk and Sharp (2004)	WCST	Judgment
Fournier-Vicente et al. (2008)	Plus–minus task	Judgment
	Number–letter task	Judgment
	Global–local	Dimension
Friedman et al. (2007, 2006, 2011)	Number–letter task	Judgment
	Color–shape	Judgment
	Category-switch	Judgment
Fuhs and Day (2011)	Flexible item selection[c]	Judgment
Hedden and Yoon (2006)	Plus–minus task	Judgment
	WCST	Judgment
	TMT	Stimulus set
Huizinga et al. (2006)	Local–global	Dimension
	Dots–triangles	Judgment
	Smiling faces	Judgment
Hull et al. (2008)	Global–local nonverbal	Dimension
	Global–local verbal	Dimension
	Plus–minus task	Judgment
	WCST	Judgment
Lee et al. (2013, 2012)	Picture–symbol task	Judgment
Lehto et al. (2003)	TMT	Stimulus set
McCabe et al. (2010)	WCST	Judgment
	Mental control[b]	Stimulus set
Miller et al. (2012)	DCCST	Judgment
Miyake et al. (2000)	Plus–minus task	Judgment
	Number–letter task	Judgment
	Global–local	Dimension
Reimers and Maylor (2005)	Gender-emotion switching	Judgment
Ropovik (2014)	TMT	Stimulus set
	WCST	Judgment
Rose et al. (2011)	TMT	Stimulus set
	Intra-/extradimensional[c]	Judgment
Salthouse (2005)	WCST	Judgment
	TMT	Stimulus set
van der Sluis et al. (2007)	Objects shifting	Judgment
	Symbol shifting	Judgment
	Place shifting	Judgment
	TMT	Stimulus set
van der Ven et al. (2012)	Animal shifting	Judgment
	TMT	Stimulus set
	Sorting task[c]	Judgment
Vaughan and Giovanello (2010)	Number–letter	Judgment
	Global–local	Dimension
	Parity–magnitude	Judgment
Willoughby et al. (2010)	Item selection task	Judgment
Willoughby et al. (2012)	Something's the same[c]	Judgment

Note. TMT = Trail-Making Test; WCST = Wisconsin Card Sorting Test; DCCST = Dimensional Change Card Sort Test.
[a] Variant of the Wisconsin Card Sorting Test. [b] Variant of the Trail Making Test. [c] Variant of the Dimensional Change Card Sort Test.

termined threshold of proportion correct responses dependent on the number of response options (75% for two-choice tasks, 66.7% for three-choice tasks, and 62.5% for four-choice tasks), resulting in a final sample size of 119 (94 female, age $M = 23.08$, $SD = 3.94$; range = 18–34 years). All participants reported normal or corrected-to-normal vision and none of them exhibited color blindness as determined by Ishihara's color test (Ishihara, 2003).

Apparatus

Participants were tested in small groups of two to five in one session of approximately 3 to 4 hr duration. Stimulus presentation and response recording were controlled by IBM-compatible PCs with a Windows 7 OS. The stimuli were presented on 19-in. color monitors (DELL 1905FP) with a

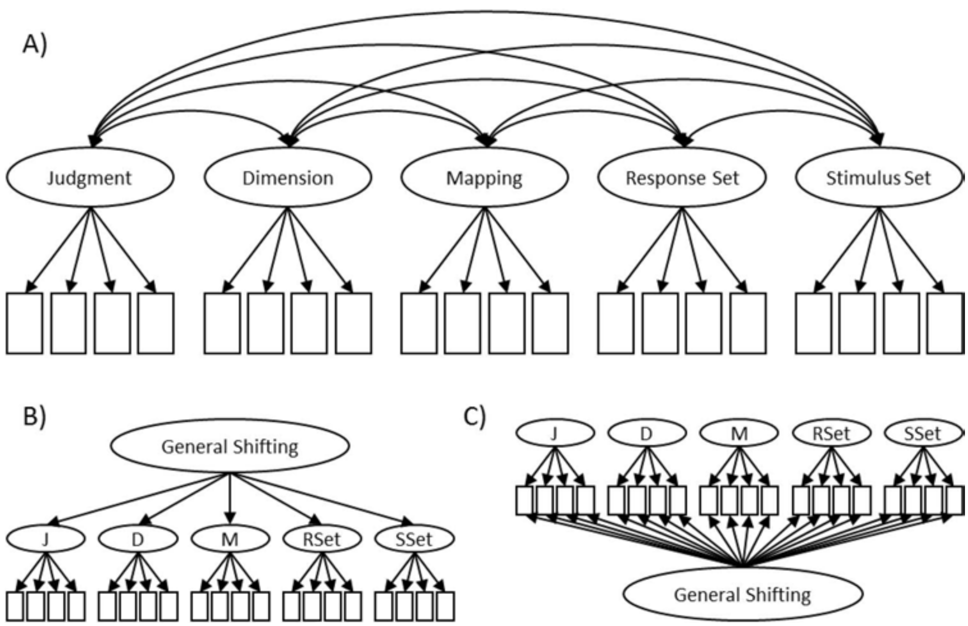


Figure 3. Baseline models for the series of models we investigated in this study. Rectangles denote manifest (observed) variables (i.e., switch costs from the respective task pairs); ellipses represent latent factors. Single-headed arrows denote linear regressions; double-headed arrows represent correlations. (A) Measurement model with the proposed five separable yet correlated shifting components. (B) Hierarchical model with a second-order general shifting factor with loadings from the latent component factors. (C) Bifactor model in which all tasks load on both a general shifting factor and on latent component-specific factors. J = judgment; D = dimension; M = mapping; RSet = response set; SSet = stimulus set.

resolution of 1280 X 1024 pixels and a refresh rate of 60 Hz. The screen was located centrally on a desk in front of the participants and viewing distance was approximately 50 cm. The experiment was programmed using the Java-based software package Tatool (von Bastian, Locher, & Ruflin, 2013). All tasks and materials are available for download on Tatool Web (http://www.tatool-web.com/#/doc/main-lib.html).

General Procedure

For each shifting component (see Figure 3A), four task pairs were used (for an overview, see Table 2). Two of these task pairs were two-choice, one a three-choice, and one a four-choice task. Furthermore, we used a cuing procedure, which allowed for a pseudorandomized task order. The cues (see Figure 4) were presented centered above the target stimuli on a white background screen. The cues appeared with a CTI of 150 ms before the target stimuli and remained on the screen during target presentation. The targets were presented centered on the screen. The target stimuli—along with the cues—remained on the screen until participants responded. After each trial, participants received immediate visual feedback (a green check mark or a red cross) about the correctness of their response. This feedback remained on screen for the full response-cue interval of 500 ms. For responding, participants had to press keys on a Swiss-German QWERTZ keyboard (see Table 2 for the details regarding each task).

For each task pair, participants performed 24 practice trials (which were not further analyzed) and 144 experimental trials. One half of the trials were repetition trials, and the other half of the trials were switch trials, with pseudorandomized order of repetition and switch trials. The task pairs for each shifting component were presented in blocks so that participants always performed all task pairs for a given component before switching to those of the next component. Across all shifting components and task pairs, two testing orders were established, that is, half of the participants performed the task pairs in the one order and the other half of participants in the reversed order (see Appendix, Table A1, for details). The task pairs within a shifting component were arranged in a way that participants always started and ended with one of the two-choice tasks, and the three- and four-choice tasks were administered in-between.

Tasks and Stimuli

An overview of all task pairs along with the description of the stimuli used is given in Table 2. *Judgment shifting* task pairs required participants to switch between varying classification rules. For example, in the three-choice task pair, participants had to determine either the color or the shape of objects presented. In *dimension shifting* task pairs, the classification rule remained the same, but participants had to switch between which of two stimulus dimensions to attend to. For example, in the Two-Choice B variant, character strings containing each a digit and a letter were presented. Participants had to attend to and identify the position of either the digit or the letter. In *mapping shifting* task pairs, participants had to switch between stimulus–response mappings. For example, in the four-choice

Table 2
Task Pairs Used to Measure the Five Proposed Shifting Components

Variant	Description	Adapted from	Response keys
	Judgment shifting		
Two-Choice A	Classify a digit according to its parity (odd or even) or magnitude (smaller or greater than 5).	Logan and Bundesen (2003)	arrow-down: even or smaller than 5; arrow-up: odd or greater than 5
Two-Choice B	Classify line drawing of an object or animal (Snodgrass and Vanderwart, 1980; Szekely et al., 2004) according to its animacy (living or non-living) or airworthiness (able to fly or not).	Mayr and Kliegl (2000)[a]	arrow-left: living or airworthy; arrow-right: nonliving or nonairworthy
Three-choice	Classify a colored shape according to its shape or color.	Mayr and Kliegl (2000)	1: circle or red; 2: triangle or green; 3: square or blue
Four-choice	Identify the position of a black dot when mentally moving it to the next corner within a rectangle either in clockwise or counterclockwise direction.	Mayr (2002)	1: bottom left; 3: bottom right; 7: top left; 9: top right
	Dimension shifting		
Two-Choice A	For digits made up of multiple small digits, classify either the large digit (global) or the small digits (local) as odd/even.	R. Hübner et al. (2001)	arrow-up: odd; arrow-down: even
Two-Choice B	For strings consisting of a digit (2 or 3) and a letter (A or B) separated by three "#" symbols (e.g., "A###2"), identify the digit's or letter's position.	Bisiacchi et al. (2009)	arrow-left: left; arrow-right: right
Three-choice	For vertically arranged combinations of a colored letter and a colored digit, identify the letter's or digit's color	Rey-Mermet and Meier (2014)	1: red; 2: green; 3: blue
Four-choice	For four rectangles arranged in a square, identify the position of the rectangle deviating in orientation or in size.	Mayr and Keele (2000)	1: bottom left; 3: bottom right; 7: top left; 9: top right
	Mapping shifting		
Two-Choice A	Classify the color of simple shapes.	Meiran (1996)	Arrow-left: green; arrow-right: red (and reverse)
Two-Choice B	Decide whether two simultaneously presented fractal images are the same or different.	von Bastian and Oberauer (2013)	arrow-left: same; arrow-right: different (and reverse)
Three-choice	Classify a word as city, river, or country.	von Bastian and Oberauer (2013)	8: city; 5: country; 2: river (and reverse)
Four-choice	Identify the mathematical symbol presented.	Duncan (1978)	v: -;-; b: x; n: -; m: + (and reverse)
	Response set shifting		
Two-Choice A	Classify a letter (A, E, O, U, B, G, R, or S) as consonant or vowel.	Rogers and Monsell (1995)	v (vowel) and b (consonant), or 1 (vowel) and 4 (consonant)
Two-Choice B	Decide whether a smiley is shown in the top or the bottom of two vertically arranged squares.	Meiran (1996, 2000a)	9 (top) and 1 (bottom), or 7 (top) and 3 (bottom)
Three-choice	Classify the type of filling of a circle.		a (black), s (grey), and d (none), or 2 (black), 5 (grey), and 8 (none)
Four-choice	Identify the color of the letter X.	Kunde (2001)	g (red), z[b] (green), h (blue), and b (yellow), or 1 (red), 7 (green), 9 (blue), and 3 (yellow)
	Stimulus set shifting		
Two-Choice A	Classify either a shape (triangle, diamond, circle, or ellipse) or a mathematical symbol (y, ±, 00, or =) as round or angular.	Druey (2014)	arrow-up: round; arrow-down: angular
Two-Choice B	For strings of five letters or symbols, decide whether the characters at the second and fourth position are the same or different.	Bisiacchi et al. (2009)	arrow-left: same; arrow-right: different
Three-choice	Identify the position of a letter (A, B, or C) or digit (1, 2, or 3) in terms of its natural order.	Karle et al. (2010)	1: first; 2: second; 3: third
Four-choice	In a display of four flowers or cars arranged in a square, identify the position of the one with a color deviating from the others.		1: bottom left; 3: bottom right; 7: top left; 9: top right

Note. All number keys used refer to the numeric key pad. Stimuli were presented individually where not noted otherwise.
[a] Only animacy task. [b] Corresponds to *Y* on English QWERTY keyboards.

1272 VON BASTIAN AND DRUEY

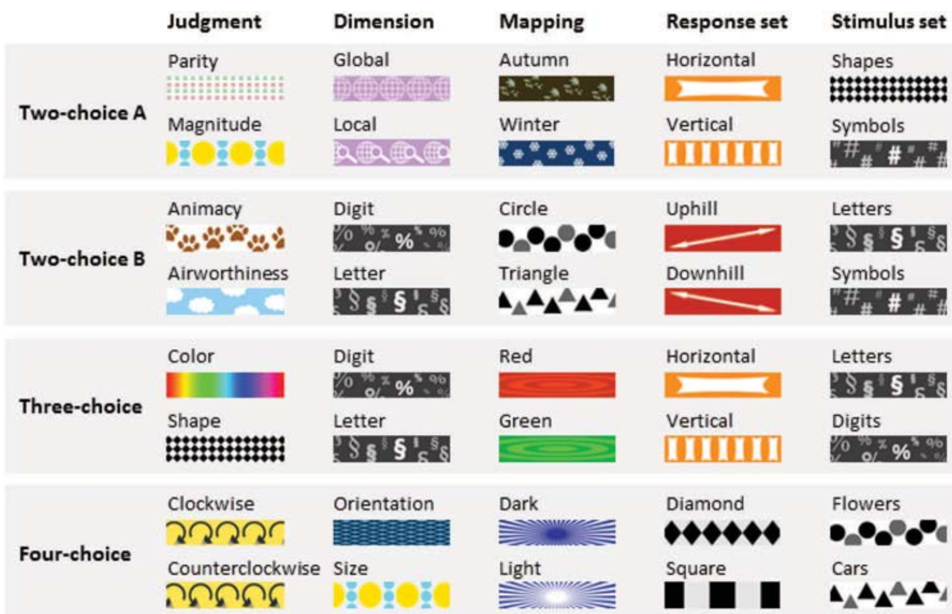


Figure 4. Visual cues indicating the currently relevant task in the shifting task pairs along with the labels used for the task instructions given to participants. See the online article for the color version of this figure.

task pair, the mathematical symbols for division ("-;-"), multiplication ("x"), subtraction ("-"), and addition ("+") were mapped to the keys *V*, *B*, *N*, and *M*, respectively. Depending on the cue, the mapping had to be reversed, so that "-;-" was no longer mapped to *V* but to *M* instead.[2] *Response set* shifting required participants to switch between two different sets of response keys. For example, in the three-choice variant, participants had to determine whether a circle was solid black, solid gray, or empty, using either a horizontal key set *A*, *S*, and *D*, or a vertical key set on the numeric pad consisting of *8*, *5*, and *2*. In the *stimulus set* shifting task pairs, the stimulus material switched depending on the cue displayed. As a consequence, different from the other shifting task pairs administered, stimulus set shifting involved univalent stimuli. For example, in the Two-Choice A task pair, participants were asked to classify either shapes (triangle, diamond, circle or ellipse) or symbols ("ɣ," "±," "ꝏ," or "=") as being round or angular.

RTs to correct responses and switch costs served as dependent measures. Switch costs were based on the log-transformed RTs to correct responses, thereby avoiding general speed or scaling effects driving the relationships between components (cf. Meiran, 1996; Ratcliff, 1993). As log-transformed switch costs can suffer from problems inherent to difference scores such as lower reliability (e.g., Cronbach & Furby, 1970; Edwards, 2001), we additionally ran all analyses with the individual studentized residuals from a linear regression model predicting the switch RTs from the repetition RTs (cf. Ecker, Lewandowsky, & Oberauer, 2014). These analyses led to the same conclusions as those based on log-transformed switch costs (see supplemental materials for the detailed results; see Lo & Andrews, 2015, for a detailed discussion on potential benefits and disadvantages of RT transformations).

Results

Data Preprocessing

RTs of wrong responses and RT outliers were excluded from the analyses. Outliers were defined as RTs three median absolute deviations away from the overall mean (Leys, Ley, Klein, Bernard, & Licata, 2013). Following these procedures, $M = 11.30\%$ of trials ($SD = 2.46\%$; range: 5.82% to 14.48% for the individual tasks) were excluded (see Appendix, Table A2, for descriptive statistics of the error rates). To eliminate unwanted variance caused by the two orders of test administration, one order was arbitrarily chosen as reference condition whereas the data from the other order was corrected by subtracting the differences in means between the two orders for each variable (cf. von Bastian & Oberauer, 2013; von Bastian, et al., 2016). All analyses were based on order-corrected RTs and error rates. The data are available on the Open Science Framework (https://osf.io/u8bh2).

Preliminary Analyses

First, we tested whether all shifting tasks elicited significant switch costs. For this purpose, we ran paired *t* tests for each task comparing RTs in switch to RTs in repetition trials. As listed in Table 3, most tasks yielded significant effects of switching, with two exceptions: the judgment switching Four-choice task, and the

[2] Piloting of the test battery showed that the three- and four-choice mapping tasks proved to be extremely difficult as indicated by error rates close to chance level. Therefore, we decided to simplify these tasks by only using reversed mapping conditions throughout (necessarily so in the two-response tasks).

Table 3
Descriptive Statistics for and Effects of Switching on Raw Response Times

Task	Switch *M*	Switch *SD*	Repetition *M*	Repetition *SD*	Effect [95% CI]	*t*(118)	*p*
			Judgment				
2A	1,161	240	838	150	323 [294, 352]	22.00	<.001
2B	910	196	773	147	137 [117, 158]	13.19	<.001
3	1,215	226	905	188	310 [280, 341]	20.38	<.001
4	1,346	350	1,334	351	12 [–9, 33]	1.15	.252
			Dimension				
2A	1,453	347	1,244	349	209 [181, 238]	14.40	<.001
2B	720	169	668	165	53 [40, 66]	8.05	<.001
3	1,582	302	1,425	305	157 [122, 192]	8.92	<.001
4	853	209	774	188	80 [63, 96]	9.36	<.001
			Mapping				
2A	1,013	225	730	142	283 [255, 310]	20.46	<.001
2B	1,135	285	917	175	218 [187, 249]	13.81	<.001
3	1,268	247	1,093	190	174 [149, 199]	13.85	<.001
4	1,311	280	1,007	222	304 [276, 332]	21.77	<.001
			Response set				
2A	1,023	223	725	113	298 [270, 326]	20.95	<.001
2B	990	252	906	218	84 [66, 102]	9.11	<.001
3	1,074	214	920	216	154 [132, 176]	13.83	<.001
4	1,339	252	1,247	295	92 [62, 122]	6.10	<.001
			Stimulus set				
2A	564	70	553	63	11 [6, 17]	4.44	<.001
2B	627	77	627	72	1 [–5, 4]	.35	.728
3	545	57	521	50	25 [20, 30]	9.78	<.001
4	552	88	543	109	10 [4, 16]	3.16	.002

Note. 2A = Two-Choice A; 2B = Two-Choice B; 3 = three-choice; 4 = four-choice.

stimulus set switching Two-Choice B task. These two tasks were therefore excluded from further analyses.[3]

Descriptive statistics and reliabilities for log-transformed switch costs are presented in Table 4 (including the excluded tasks for completeness). Most included tasks showed acceptable reliabilities ($\alpha \geq .66$), except for mapping three-choice ($\alpha = .48$), and stimulus set Two-Choice A ($\alpha = .40$). In addition, most between-tasks zero-order correlations (see Table 5) within each proposed shifting type were significantly positive, with only one exception (response set Two-Choice B). Performance in this task also showed no significant positive correlation with any other task (and also did not load significantly on the later modeled response set factor). As we realized only post hoc, the response set Two-Choice B task pair differed qualitatively from all other task pairs in this component in that it was the only one involving spatially overlapping response sets. Given these considerations, we decided to exclude this task from further analyses. However, including this task yielded qualitatively the same results and led to the same conclusions.[4]

Structural Analyses

To examine the structure of shifting factors, we ran latent-variable confirmatory factor analyses (CFA) using the Lavaan package (Version 0.5–19, Rosseel, 2012) in R (Version 3.2.2, R Core Team, 2014). Model fit was evaluated examining the chi-square statistic (χ^2), the root mean square error of approximation (RMSEA) alongside its 90% CI and the standardized root mean squared residual (SRMR). Conventional standards indicating good fit are values smaller than .06 for RMSEA and values less than .08 for the SRMR (Hu & Bentler, 1999).

It would be desirable to additionally report the comparative fit index (CFI). The CFI compares the fit of the null model in which all covariances are fixed to zero (reflecting no communalities across tasks) to the fit of the current model. However, cases where the RMSEA of the null model is lower than .158 render the CFI noninterpretable (Kenny, 2015). We therefore examined the RMSEA of the null model using the semTools package (Version 0.4–14,

[3] A tentative explanation for the absence of switch costs in the Four-Choice judgment switching task is that it was the only task that required an additional mental transformation of the stimulus displayed (i.e., mental rotation). Hence, the greater difficulty of this task may have obscured switch costs. As we administered a considerably lower number of (practice) trials than was the case in the original study (Mayr, 2002), it is possible that more trials would have allowed for switch costs to become manifest.

[4] For Model 1 to converge when including this task pair, its loading on the response set factor had to be fixed to zero, reflecting its low correlation with the other response set task pairs.

Table 4
Descriptive Statistics and Reliabilities for Log-Transformed Switch Costs

Task	*M*	*SD*	Min	Max	Skewness	Kurtosis	Reliability[a]
			Judgment				
2A	.31	.14	.60	.11	.20	.13	.88
2B	.15	.11	.50	.07	.73	.48	.86
3	.29	.13	.73	.01	.09	.07	.79
4	.01	.08	.23	.22	.11	.37	.41
			Dimension				
2A	.16	.11	.51	.10	.32	.29	.70
2B	.07	.08	.27	.20	.02	.34	.79
3	.12	.11	.48	.16	.34	.19	.75
4	.09	.09	.32	.19	.02	.30	.78
			Mapping				
2A	.32	.13	.72	.00	.51	.75	.86
2B	.20	.11	.51	.02	.54	.02	.81
3	.14	.08	.52	.03	1.11	2.63	.48
4	.25	.10	.54	.01	.11	.03	.66
			Response set				
2A	.32	.12	.67	.09	.70	.37	.89
2B	.08	.09	.35	.12	.51	.50	.73
3	.16	.11	.42	.18	.38	.02	.78
4	.09	.11	.33	.27	.43	.25	.69
			Stimulus set				
2A	.02	.04	.15	.08	.20	.02	.40
2B	.00	.04	.09	.10	.07	.34	.40
3	.04	.05	.18	.08	.21	.14	.73
4	.02	.06	.11	.23	1.62	3.61	.86

Note. 2A = Two-Choice A; 2B = Two-Choice B; 3 = three-choice; 4 = four-choice; Min = minimum; Max = maximum.
[a] Split-half reliability corrected using Spearman-Brown's prophecy formula.

Jorgensen et al., 2016) and found $RMSEA_{null} = .124$.[5] Therefore, we refrained from reporting the CFI or other incremental fit indices. Nested models were compared by change in model fit assessed with chi-square difference tests.

To establish a baseline model, we first evaluated the fit of the a priori theoretical model (consisting of five separate but correlated latent shifting factors) with the observed data. We then compared the fit of multiple alternative models to this baseline model. In the first series of comparisons, we examined the factors' separability by reducing the number of latent factors toward full unity (i.e., a single factor), and by testing full diversity (i.e., five uncorrelated factors). Next, we investigated the extent to which unity and diversity may both be present in the shifting structure by evaluating a series of hierarchical and bifactor models.

Theoretical baseline model. The proposed theoretical model with five correlated factors for judgment, dimension, mapping, response set, and stimulus set shifting resulted in $\chi^2(110) = 136.50$, $p = .044$, RMSEA = .045 [.008; .068], SRMR = .072 (see also Table 6). Hence, the model fit the data well according to the RMSEA and SRMR. As illustrated in Figure 5, all tasks loaded significantly on their respective factor, except the stimulus set Two-Choice A task. This task yielded a nonsignificant negative residual variance, which therefore had to be fixed to zero.[6] Furthermore, correlations between the latent factors were generally significantly positive, except for the stimulus set factor, which was solely related to mapping. Finally, all latent factors exhibited significant variance, except for the dimension factor ($p = .063$).[7]

Separability of Shifting Factors

After having established the baseline model illustrated in Figure 5, we examined whether fewer factors were sufficient to reflect the structure of shifting. For this purpose, we combined the factors ordered by the size of their covariance. We started with the mapping factor, because switching of the exact stimulus–response mappings is an inherent feature across all types of the included shifting tasks (except those representing the dimension shifting component), and combined it with the response set factor (Model 2a in Table 6). However, Model 2a's fit indices were overall worse, $\chi^2(114) = 149.15$, $p = .015$, RMSEA = .051 [.024; .072], SRMR = .075, with Model 1 fitting the data significantly better, $\Delta\chi^2(4) = 12.65$, $p = .013$. Therefore, we did not expect to find that the following models combining mapping, response set, and judgment (Model 2b), mapping, response set, judgment, and dimension (Model 2c), or the full unity model (Model 2d[8]) would yield better fit, but listed them for completeness in Table 6. Similarly, full diversity (i.e., five uncorrelated factors, Model 3) resulted in a poor fit. Taken together, based on this series of model comparisons, the theoretical baseline model resulted in the best fit. In this model the five factors were significantly correlated, except

[5] Note that the general fit of the null model was still poor, with $\chi^2(136) = 384.65$, $p < .001$, and significantly worse than that of the proposed theoretical baseline model, $\Delta\chi^2(26) = 248.14$, $p < .001$.

[6] The negative residual variance was possibly caused by its low reliability or by this task correlating moderately with both the three-choice ($r = .30$) and four-choice variant ($r = .27$), whereas the latter two tasks were unrelated ($r = .03$). This pattern may reflect a larger overlap in required processes for the Two-Choice A and the three-choice variant, but a larger overlap in material domain between the Two-Choice A and the four-choice variant. More specifically, the Two-Choice A and the three-choice stimulus set shifting tasks require participants to apply an additional (but, in contrast to the Judgment switching tasks, constant) semantic classification rule (i.e., the contour or the natural order, respectively), whereas perception-based identification of the deviant item is sufficient in the four-choice task. However, both the Two-Choice A and the four-choice tasks use nonverbal stimulus sets, whereas the three-choice task uses verbal–numerical stimulus sets. The negative residual variance was eliminated by fixing it to zero.

[7] The zero-order correlations between task pairs (see Table 5) suggest that the shared variance between the dimension task pairs may have turned out nonsignificant due to the very low correlation between the Two-Choice A and the Two-Choice B task. Indeed, exploratory analyses, in which each task was stepwise removed from the model, showed that the variance of the dimension factor was significant when either of the task pairs had been excluded, with slightly better fit without the Two-Choice A task pair. All analyses reported were therefore also performed without the Two-Choice A task, which yielded largely identical results with only a few exceptions. First, fit indices revealed generally better fit when excluding this task pair (e.g., for Model 1: $\chi^2(95) = 111.21$, $p = .123$, RMSEA = .038 [.000; .064], SRMR = .068). Second, the correlation between judgment and dimension shifting dropped to .27 and was nonsignificant ($p \geq .066$) in Models 1 and 2a. The results of the model comparisons, however, were the same independent of whether the Two-Choice A task pair was excluded or not.

[8] As all tasks were loading on a single factor, the previously fixed residual variance for the Two-Choice A set shifting task could be set to vary freely.

Table 5
Between-Tasks Correlations of Log-Transformed Switch Costs

Task	Judgment 2A	Judgment 2B	Judgment 3	Dimension 2A	Dimension 2B	Dimension 3	Dimension 4	Mapping 2A	Mapping 2B	Mapping 3	Mapping 4	Response 2A	Response 2B	Response 3	Response 4	Stimulus set 2A	Stimulus set 3
Judgment																	
2B	**.40**																
3	**.40**	**.22**															
Dimension																	
2A	**.19**	**.26**	**.23**														
2B	.02	.00	**.23**	.02													
3	.03	.02	**.23**	**.32**	**.29**												
4	**.20**	.09	.16	.17	**.32**	**.28**											
Mapping																	
2A	**.18**	**.19**	.11	**.26**	.15	**.19**	**.23**										
2B	**.27**	**.32**	**.27**	.11	.05	**.24**	.07	**.20**									
3	.03	.17	.08	**.21**	.16	**.25**	**.20**	.15	.18								
4	**.32**	**.23**	**.19**	.13	.11	**.20**	.13	**.26**	**.33**	**.27**							
Response set																	
2A	**.43**	**.29**	**.33**	.12	.06	.18	**.20**	**.24**	**.31**	.17	**.38**						
2B	.02	.06	.03	**–.19**	.07	.01	.01	.01	.04	.01	.02	.00					
3	.09	**.29**	.09	.16	.16	.09	.06	**.22**	**.23**	.11	**.21**	**.30**	.05				
4	**.19**	.14	.00	.06	.14	.13	.08	.14	.12	.04	**.24**	.10	.07	**.38**			
Stimulus set																	
2A	.03	.01	.03	.13	.09	.08	.03	.23	.16	.06	.15	.02	.18	.10	.07		
3	.16	.04	.05	.05	.11	.00	.02	.16	.01	.14	.09	.11	.00	.00	**.21**	**.30**	
4	.09	.08	.12	.10	.04	.02	.04	.05	.01	.06	.01	.06	**–.27**	.07	.01	**.27**	.03

Note. Significant correlations ($p < .05$) are printed in bold. 2A = Two-Choice A; 2B = Two-Choice B; 3 = three-choice; 4 = four-choice.

for Stimulus Set shifting, which was solely related to mapping shifting.

Unity and Diversity of Shifting

Given that four of the five factors correlated at least moderately well, we tested whether an underlying common shifting process could be captured by a general shifting factor (gShifting).

Second-order factor model. In this series of models, we conceptualized gShifting as hierarchical second-order factor with loadings from all five factors. The first model (Model 4a) resulted in an acceptable fit, $\chi^2(115) = 150.41$, $p = .015$, RMSEA = .051 [.024; .072], SRMR = .075. As Mapping loaded on gShifting with .99, we set the Mapping tasks to load directly on gShifting, resulting in Model 4b. To account for the observation that Stimulus Set shifting was only correlated with Mapping but none of the other factors, we modeled Stimulus Set shifting to covary with instead of loading on gShifting (Model 4c), resulting in a mathematically identical model. The fit of Models 4b/c was acceptable, $\chi^2(116) = 150.41$, $p = .017$, RMSEA = .050 [.022; .071], SRMR = .075, and statistically indistinguishable from Model 4a's fit, $\Delta\chi^2(1) < 0.01$, $p = .932$. Hence, Model 4b/c was retained. Whereas the loadings of both the judgment and the dimension shifting factors on gShifting were of a medium size and resulted in significant residual variance, the response set shifting factor's loading on gShifting was high (.91) and its residual variance nonsignificant ($p = .416$). Therefore, we tested whether a separate response set factor was appropriate by setting the response set tasks to load directly on gShifting. The fit of this Model 4d was again acceptable, $\chi^2(117) = 151.02$, $p = .019$, RMSEA = .049 [.021; .071], SRMR = .075, and not significantly different from the fit of Model 4b/c, $\Delta\chi^2(1) = 0.61$, $p = .435$. Thus, as the fuller Model 4b/c did not yield significantly better fit than the more restricted Model 4d, the latter was retained (see Figure 6). Next, we compared Model 4d with a Model 4e in which the Judgment shifting tasks were also set to load directly on the gShifting factor. Although the model fit was still acceptable, $\chi^2(118) = 159.04$, $p = .007$, RMSEA = .054 [.029; .075], SRMR = .078, it was significantly worse than for Model 4d, $\Delta\chi^2(1) = 8.01$, $p = .005$. Hence, in this series of second-order factor models, Model 4d was the best fitting one. However, its fit was still significantly worse than that of the theoretical baseline model (Model 1), $\Delta\chi^2(7) = 14.52$, $p = .043$.

Bifactor model. Model 4d can also be conceptualized as a bifactor model, with all tasks (except for the stimulus-set shifting tasks) loading on gShifting, and the judgment and dimension shifting tasks additionally loading on judgment-specific and dimension-specific factors (for similar approaches, see Friedman et al., 2008; van der Sluis, et al., 2007). Model 5a (see Figure 7) fitted the data similarly well, $\chi^2(113) = 146.00$, $p = .020$, RMSEA = .050 [.021; .071], SRMR = .073, with no significant difference in fit in comparison to Model 4d, $\Delta\chi^2(4) = 5.03$, $p = .285$. Figure 7 illustrates that despite the

Table 6
Fit Statistics for the Confirmatory Factor Analysis Models

Model	Description	*df*	χ^2	*p*	RMSEA [90% CI]	SRMR
1	Five correlated factors	110	136.50	.044	.045 [.008; .068]	.072
2a	Four correlated factors, Mapping = Response Set	114	149.15	.015	.051 [.024; .072]	.075
2b	Three correlated factors, Mapping = Response Set = Rule	117	159.01	.006	.055 [.031; .075]	.077
2c	Two correlated factors, Mapping = Response Set = Rule = Dimension	119	175.72	.001	.063 [.042; .082]	.081
2d	Unity (single factor)[a]	119	193.05	<.001	.072 [.053; .091]	.085
3	Full diversity (five orthogonal factors)[b]	121	222.71	<.001	.084 [.066; .101]	.143
4a	Second-order factor with five first-order factors	115	150.41	.015	.051 [.024; .072]	.075
4b/4c	Second-order factor with four first-order factors	116	150.41	.017	.050 [.022; .071]	.075
4d	Second-order factor with three first-order factors	117	151.02	.019	.049 [.021; .071]	.075
4e	Second-order factor with two first-order factors	118	159.04	.007	.054 [.029; .075]	.078
5a	Common shifting factor with two uncorrelated specific factors[c]	113	146.00	.020	.050 [.021; .071]	.073
5b	Common shifting factor with one uncorrelated specific factor[c]	117	164.01	.003	.058 [.035; .078]	.078

Note. The comparative fit index is not reported as it cannot be interpreted because of null model's root mean square error of approximation (RMSEA) of <.158.
[a] Different than for the other models, the residual variance for the Two-Choice A set shifting task was set to vary freely. [b] Fitting resulted in a negative (but nonsignificant) residual variance for the three-choice response set shifting task, which was therefore fixed to zero. [c] Fitting resulted in a negative (but nonsignificant) residual variance for the Two-Choice A judgment shifting task, which was therefore fixed to zero.

reasonable fit, the judgment-specific factor was mainly defined by the judgment Two-Choice A shifting task, the variance of which had to be fixed to zero. Furthermore, the dimension shifting tasks did not load equally well on the gShifting and the dimension-specific factor. Therefore, we additionally explored Model 5b omitting the dimension-specific factor which, however, resulted in significantly poorer fit than Model 5a, ilx^2(4) = 18.01, $p = .001$, and was therefore discarded. Next, Model 5a was compared with the theoretical baseline model, showing that the bifactor model fit significantly worse than the correlational model, ilx^2(3) = 9.49, $p = .023$.

Alternative Models

We also tested alternative factor structures in which the tasks were grouped by the number of response options (i.e., consisting of factors two-choice, three-choice, and four-choice tasks) or by material (i.e., verbal and nonverbal materials). However, both factor models yielded inadmissible solutions due to inseparable factors (i.e., the three-choice factor was inseparable from the two-choice and the four-choice factor, as were the two material factors). We therefore had to discard these models.

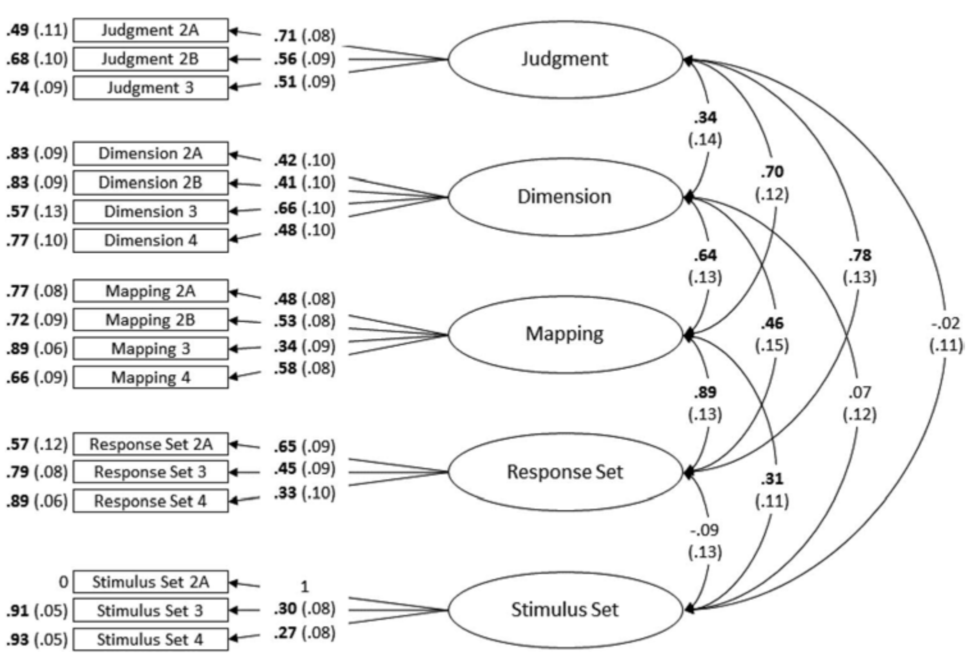


Figure 5. Five-factor measurement model of shifting (Model 1). Bold numbers indicate significance ($p < .05$), standard errors are given in parentheses. All latent factor variances were significant ($p < .05$) except for the dimension factor ($p = .063$).

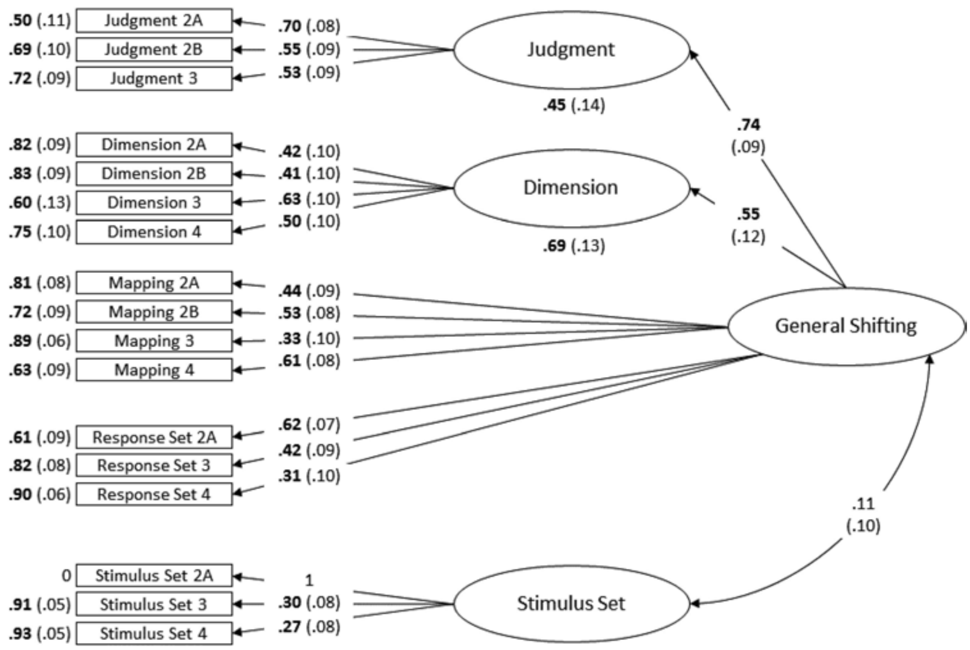


Figure 6. Hierarchical model with a general shifting second-order factor and loadings from two first-order factors (Model 4d). Bold numbers indicate significance ($p < .05$), standard errors are given in parentheses. Variances of general shifting and stimulus set shifting were significant ($p < .030$).

Final Model

In summary, Model 1 (see Figure 5) with five correlated factors presented the best fit to the data. In this model, shifting between mappings showed most commonality with shifting between response sets. Despite being relatively strongly related (coefficient estimate = .89), nested model comparisons suggested that mapping and response set still represented separable factors. Whereas shifting between sets correlated only with shifting between mappings, all other types of shifting were at least moderately interrelated, with coefficient estimates ranging between .34 and .89.

Discussion

In this study, we took an individual differences approach to investigate the ability of shifting between mental sets. On the basis of an extension of Kleinsorge's structural-hierarchical model of task set (2004; Kleinsorge & Heuer, 1999), we examined whether the 20 task pairs we adapted from the task-switch literature loaded on the five proposed components of mental set shifting (i.e., judgment, dimension, mapping, response set, and stimulus set shifting). Moreover, we examined how these components were interrelated on the latent-factor level. Finally, we examined whether a common factor could account for considerable shared variance between the five components, reflecting a general underlying shifting ability.

Key Findings

Most of the tasks we selected indeed loaded on the proposed shifting components we assigned them to. Moreover, the majority of task pairs produced significant switch costs with relatively high reliabilities, thereby providing a good basis for modeling the factorial structure of shifting. Model comparisons showed that four of the five components were closely related but still separable. The fifth component, stimulus set shifting, was related to the mapping component only. Regarding the question whether there is a general shifting factor reflecting a common underlying ability, the answer is more complex. In the best fitting models of these series, the mapping and response set tasks loaded directly on the gShifting factor, with additional factors being required for judgment, dimension, and stimulus set shifting tasks. Note, though, that these models, despite providing reasonable accounts for the data, nevertheless fit worse than the baseline model, thereby justifying the omission of a common gShifting factor.

In the following, we consider the results of our study and their implications in more detail. These concern theoretical implications, limitations, and recommendations for future investigations of shifting abilities.

Theoretical Implications

One crucial aspect of our results is the central role of the Mapping component. It was strongly correlated with every other component and indistinguishable from the higher-order gShifting factor in the hierarchical and bifactor models. From a purely methodological viewpoint, this finding is not that surprising, as shifts of the SR-mappings were required for the tasks in all components in case of a task switch (except for dimension shifting, see Figure 2). Conceptually, though, the hierarchy of shifting components resulting from our individual differences approach and the hierarchy of the task-set features in the framework of Kleinsorge and Heuer (1999; Kleinsorge, 2004) reflect two differ-

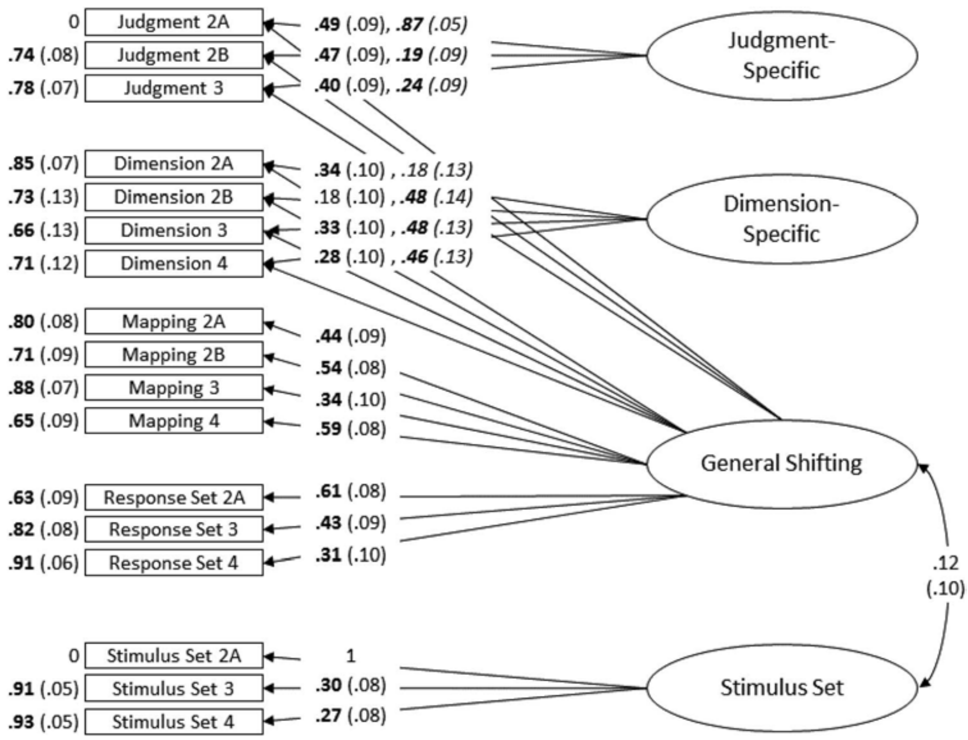


Figure 7. Bifactor model with a general shifting factor and two uncorrelated factors to reflect specific variance for judgment and dimension shifting (Model 5a). Bold numbers indicate significance ($p < .05$), standard errors are given in parentheses. Numbers in italics refer to loadings (alongside their respective standard errors) on the specific factors. All latent factor variances were significant ($p < .05$) except for the dimension factor ($p = .485$).

ent perspectives on the same phenomenon. Whereas Kleinsorge and Heuer (1999; Kleinsorge, 2004) took a functional perspective guided by the order of selection processes in a task-switch situation, we took a structural perspective, focusing on individual differences in the shifting components and their relationships. From a functional viewpoint, the judgment is selected first, which then determines the relevant stimulus dimension(s) and the (set of) relevant mappings. Last, the response is selected based on the selected mappings. From our structural perspective, the mapping component is the most central one, reflecting that mapping switches are immanent to most of the shifting components. Despite these differences in perspective, the shifting components derived from the framework of Kleinsorge (2004) and Kleinsorge and Heuer (1999) still reflected essential aspects of mental set shifting in our individual differences approach.

A second important conceptual implication pertains to the good fit of our results with Oberauer's (2009) proposed theoretical framework of procedural working memory, in which the ad hoc mapping and unmapping of input (i.e., a stimulus) to output (i.e., a response) information is essential for goal-directed and flexible cognition. More specifically, Oberauer (2009) suggested that the switch costs observed in task-switch studies are mainly due to the required switching back and forth between different sets of SR-mappings in working memory, thereby rendering these mappings the core feature of task sets. Our results support the notion that mapping shifting reflects an essential aspect of the general shifting ability.

Furthermore, we found that the response set shifting component is also highly related to all other components with the exception of stimulus set shifting. One important aspect to note when comparing our response set shifting component to the response task feature of Kleinsorge and Heuer (1999; Kleinsorge, 2004) is that there is an essential conceptual difference: Whereas Kleinsorge and Heuer (1999; Kleinsorge, 2004) conceptualize the response level primarily in terms of the individual motor responses, we are dealing with response sets, that is, chunks of responses belonging to or even making up individual tasks. Future research will have to clarify how the different response-level conceptualizations are related with respect to the hierarchical organization of task sets.

In addition, two of our findings may have implications for the processes involved in mental set shifting (i.e., the functional task-switch perspective): the relatively weak relationship between the judgment and dimension components and the isolated role of stimulus set shifting. First, regarding the relationship between the judgment and dimension components, although switching between dimensions is immanent in all of the judgment component's task pairs, the dimension component's correlation with the Judgment component was significant but unexpectedly low. This suggests that separate processes may be involved in the selection of judgments and stimulus dimensions, an explanation in line with a previous study by R. Hübner, Futterer, and Steinhauser (2001) that revealed additive effects for concurrent judgment and dimension switches. However, some caution is required when relating the two results. Whereas we focused on trial-by-trial switch costs, R.

Hübner et al. (2001) focused on mixing costs (i.e., the costs arising when comparing performance for a given task in the repetition trials of switch blocks with the performance of that same task in nonswitch blocks). Thus, it has yet to be confirmed whether similar additive effects would also arise in switch costs before final conclusions can be drawn about the functional nature of the relation between judgment and dimension shifting.

Second, stimulus set shifting clearly differed from all the other components. It only correlated with the mapping component, which likely resulted from the stimulus set task pairs still requiring shifting the exact stimulus–response associations. Consistent with previous findings from the literature on switching between tasks using univalent stimuli (e.g., Meiran, 2000b; Ruthruff, Remington, & Johnston, 2001), we observed only small but significant and relatively reliable switch costs for three out of four Stimulus Set task pairs. At the same time, stimulus set shifting was related only weakly to a general shifting ability in the hierarchical and bifactor models. Therefore, the question arising here is what actually causes the costs under these conditions in the respective tasks. One possible explanation focuses on differences between visual-spatial and executive attentional processes. Miyake and colleagues (2000) noted that pure shifts of visual-spatial attention (i.e., alerting and orienting responses) are to be distinguished from "executive-oriented shifts" (p. 56), which involve different attentional brain networks (cf. Fan, McCandliss, Fossella, Flombaum, & Posner, 2005). In stimulus set shifting tasks, switches of peripheral visual attention seem to be sufficient to successfully perform the tasks, and these shifts clearly produce much smaller costs than shifts of executive attention. This is most obvious when considering the role of the cues in these tasks. To perform well in stimulus set shifting, the cues can be completely ignored, which is not possible for the task pairs in the other components.

Limitations and Future Directions

Although we think that the current study provides an important contribution to a better understanding of the mental set shifting ability, we assessed only shifting ability without considering its relation to other cognitive abilities such as other executive functions. However, the present study can serve as a basis for future research putting shifting ability into the wider context of cognitive performance. More specifically, shifting has often been found to show different or even opposing patterns of correlations with other cognitive abilities (e.g., fluid intelligence) than other but supposedly related executive functions such as working memory (e.g., Oberauer, Süß, Wilhelm, & Wittmann, 2008; but see Draheim, Hicks, & Engle, 2016, for a possible methodological explanation), or inhibition and common executive function (cf. Miyake & Friedman, 2012). Distinguishing between subcomponents of shifting when investigating cognitive individual differences may reveal differential patterns for the components, and hence provide a novel perspective on previous findings. Furthermore, such an approach could help determining which aspects of shifting constitute it as an ability separable from other abilities, and which components are inseparable from general speed (cf. Jewsbury, Bowden, & Strauss, 2016).

Two further limitations to be considered relate to the chosen basic paradigm: For all tasks and components in this study, we relied on cued task switching. In future research, it will be necessary to show that the structure of shifting components as revealed in our study holds also for other switching paradigms such as alternating runs (Rogers & Monsell, 1995) or voluntary task switching (Arrington & Logan, 2004, 2005). Moreover, to keep the study within reasonable time limits and because it was the first attempt to disentangle different shifting components, we only used a short CTI throughout. As different processes can be predominant when participants have time to prepare for an upcoming task switch (e.g., Meiran, 1996), future studies should compare the component structure for short and long CTI conditions.

Finally, the sample size in the present study can be viewed as relatively modest. Simulation studies have shown that the sample size required for structural equation modeling depends on multiple factors, including the number of latent factors, the number of indicators, and the magnitude of factor loadings and correlations (e.g., Wolf, Harrington, Clark, & Miller, 2013). Hence, although we attempted to model five latent factors, which increases sample size demands compared with models with fewer factors, the use of up to four tasks to indicate each factor in turn decreases the required sample size. Moreover, shifting factor loadings reported in previous studies with younger adults were typically at least moderate (e.g., average loading of .66 in von Bastian & Oberauer, 2013), and correlations between factors measuring highly similar constructs can be expected to be high (e.g., .80 to .92 among three working memory factors in Wilhelm, Hildebrandt, & Oberauer, 2013). Accordingly, a Monte Carlo simulation using the simsem package (Version 0.5–14, Pornprasertmanit, Miller, Schoemann, Quick, & Jorgensen, 2016) with 10,000 iterations replicated in two artificial data sets showed that $N = 120$ yielded sufficient power (i.e., ≥80%) for a five-factor model with average factor loadings of .60 and factor correlations of .80. Furthermore, Monte Carlo simulations run for the observed model revealed that power was still ≥80% for most estimates, except for the weak correlations between stimulus set shifting and the other components. Taken together, although the modest sample size in the present study appears to be more or less unproblematic for the overall pattern of results, further research with more statistical power is warranted to investigate the role of Stimulus Set shifting in more detail.

Recommendations

Our results are informative not only with respect to the structural components of mental set shifting, but they also provide a perspective on previous and future research concerning the assessment of shifting abilities. As laid out in the Introduction, in most previous studies tasks have been used that mainly reflect Judgment shifting. In our view, this is not the best measure as it possibly covers at least two components simultaneously: Judgment-specific shifting processes and general shifting processes. On the basis of our results, for a more general examination of the shifting ability, we would suggest rather using mapping shifting tasks as they strongly correlate with all other shifting components and load most strongly on the gShifting factor in the hierarchical and bifactor models. Although, in principle, the same also holds for response set shifting, we would not recommend using response set shifting tasks as the correlations were generally lower than for the mapping shifting tasks.

Finally, the results of the present study give rise to skepticism toward the frequent use of the TMT to assess shifting abilities, as it primarily resembles stimulus set shifting. However, this component was least well correlated with the other components and probably measures shifting visual-spatial attention rather than executive shifting ability.

Conclusion

In sum, our results demonstrate that judgment, dimension, mapping, and response set shifting reflect interrelated yet separable aspects of shifting ability. Stimulus set shifting, however, is only related to mapping shifting, indicating that stimulus set shifting reflects rather specific aspects of shifting that are only little related to a general mental task set shifting ability. The mapping component was most strongly interconnected and was indistinguishable from a general second-order shifting factor. Therefore, we conclude that mapping shifting reflects the most general measure of mental set shifting and, thus, constitutes the most adequate component for future research focusing on set shifting abilities.

References

Adrover-Roig, D., Sesé, A., Barceló, F., & Palmer, A. (2012). A latent variable approach to executive control in healthy ageing. *Brain and Cognition, 78*, 284–299. http://dx.doi.org/10.1016/j.bandc.2012.01.005

Agostino, A., Johnson, J., & Pascual-Leone, J. (2010). Executive functions underlying multiplicative reasoning: Problem type matters. *Journal of Experimental Child Psychology, 105*, 286–305. http://dx.doi.org/10.1016/j.jecp.2009.09.006

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (Vol. 15, pp. 421–452). Cambridge, MA: MIT Press.

Altmann, E. M. (2004). Advance preparation in task switching: What work is being done? *Psychological Science, 15*, 616–622. http://dx.doi.org/10.1111/j.0956-7976.2004.00729.x

Altmann, E. M., & Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review, 115*, 602–639. http://dx.doi.org/10.1037/0033-295X.115.3.602

Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review, 16*, 17–42. http://dx.doi.org/10.1007/s11065-006-9002-x

Arrington, C. M., & Logan, G. D. (2004). Episodic and semantic components of the compound-stimulus strategy in the explicit task-cuing procedure. *Memory & Cognition, 32*, 965–978. http://dx.doi.org/10.3758/BF03196874

Arrington, C. M., & Logan, G. D. (2005). Voluntary task switching: Chasing the elusive homunculus. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 683–702. http://dx.doi.org/10.1037/0278-7393.31.4.683

Bisiacchi, P. S., Schiff, S., Ciccola, A., & Kliegel, M. (2009). The role of dual-task and task-switch in prospective memory: Behavioural data and neural correlates. *Neuropsychologia, 47*, 1362–1373. http://dx.doi.org/10.1016/j.neuropsychologia.2009.01.034

Chevalier, N., Dauvier, B., & Blaye, A. (2017). From prioritizing objects to prioritizing cues: A developmental shift for cognitive control. *Developmental Science*. Advance online publication. http://dx.doi.org/10.1111/desc.12534

Cronbach, L. J., & Furby, L. (1970). How we should measure "change"—or should we? *Psychological Bulletin, 74*, 68–80. http://dx.doi.org/10.1037/h0029382

de Frias, C. M., Dixon, R. A., & Strauss, E. (2009). Characterizing executive functioning in older special populations: From cognitively elite to cognitively impaired. *Neuropsychology, 23*, 778–791. http://dx.doi.org/10.1037/a0016743

Draheim, C., Hicks, K. L., & Engle, R. W. (2016). Combining reaction time and accuracy: The relationship between working memory capacity and task switching as a case example. *Perspectives on Psychological Science, 11*, 133–155. http://dx.doi.org/10.1177/1745691615596990

Dreisbach, G., Haider, H., & Kluwe, R. H. (2002). Preparatory processes in the task-switching paradigm: Evidence from the use of probability cues. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 468–483. http://dx.doi.org/10.1037/0278-7393.28.3.468

Druey, M. D. (2014). Stimulus-category and response-repetition effects in task switching: An evaluation of four explanations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 125–146. http://dx.doi.org/10.1037/a0033868

Duncan, J. (1978). Response selection in spatial choice reaction: Further evidence against associative models. *The Quarterly Journal of Experimental Psychology, 30*, 429–440. http://dx.doi.org/10.1080/00335557843000034

Ecker, U. K. H., Lewandowsky, S., & Oberauer, K. (2014). Removal of information from working memory: A specific updating process. *Journal of Memory and Language, 74*, 77–90. http://dx.doi.org/10.1016/j.jml.2013.09.003

Edwards, J. R. (2001). Ten difference score myths. *Organizational Research Methods, 4*, 265–287. http://dx.doi.org/10.1177/109442810143005

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *NeuroImage, 26*, 471–479. http://dx.doi.org/10.1016/j.neuroimage.2005.02.004

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology, 26*, 874–890. http://dx.doi.org/10.1080/13803390490510680

Fournier-Vicente, S., Larigauderie, P., & Gaonac'h, D. (2008). More dissociations and interactions within central executive functioning: A comprehensive latent-variable analysis. *Acta Psychologica, 129*, 32–48. http://dx.doi.org/10.1016/j.actpsy.2008.04.004

Friedman, N. P., Haberstick, B. C., Willcutt, E. G., Miyake, A., Young, S. E., Corley, R. P., & Hewitt, J. K. (2007). Greater attention problems during childhood predict poorer executive functioning in late adolescence. *Psychological Science, 18*, 893–900. http://dx.doi.org/10.1111/j.1467-9280.2007.01997.x

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*, 172–179. http://dx.doi.org/10.1111/j.1467-9280.2006.01681.x

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology, 47*, 1410–1430. http://dx.doi.org/10.1037/a0023750

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*, 201–225. http://dx.doi.org/10.1037/0096-3445.137.2.201

Fuhs, M. W., & Day, J. D. (2011). Verbal ability and executive functioning development in preschoolers at head start. *Developmental Psychology, 47*, 404–416. http://dx.doi.org/10.1037/a0021065

Gilbert, S. J., & Shallice, T. (2002). Task switching: A PDP model. *Cognitive Psychology, 44*, 297–337. http://dx.doi.org/10.1006/cogp.2001.0770

Hedden, T., & Yoon, C. (2006). Individual differences in executive processing predict susceptibility to interference in verbal working memory. *Neuropsychology, 20*, 511–528. http://dx.doi.org/10.1037/0894-4105.20.5.511

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. http://dx.doi.org/10.1080/10705519909540118

Hübner, M., Kluwe, R. H., Luna-Rodriguez, A., & Peters, A. (2004). Response selection difficulty and asymmetrical costs of switching between tasks and stimuli: No evidence for an exogenous component of task-set reconfiguration. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 1043–1063. http://dx.doi.org/10.1037/0096-1523.30.6.1043

Hübner, R., Futterer, T., & Steinhauser, M. (2001). On attentional control as a source of residual shift costs: Evidence from two-component task shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 640–653. http://dx.doi.org/10.1037/0278-7393.27.3.640

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44,* 2017–2036. http://dx.doi.org/10.1016/j.neuropsychologia.2006.01.010

Hull, R., Martin, R. C., Beier, M. E., Lane, D., & Hamilton, A. C. (2008). Executive function in older adults: A structural equation modeling approach. *Neuropsychology, 22,* 508–522. http://dx.doi.org/10.1037/0894-4105.22.4.508

Ishihara, S. (2003). *Ishihara's tests for colour-blindness*. Tokyo, Japan: Kanehara and Co.

Jersild, A. T. (1927). Mental set and shift. *Archives de Psychologie, 14,* 1–82.

Jewsbury, P. A., Bowden, S. C., & Strauss, M. E. (2016). Integrating the switching, inhibition, and updating model of executive function with the Cattell-Horn-Carroll model. *Journal of Experimental Psychology: General, 145,* 220–245. http://dx.doi.org/10.1037/xge0000119

Jorgensen, T. D., Pornprasertmanit, S., Miller, P., Schoemann, A., Rosseel, Y., Quick, C., . . . Longo, Y. (2016). *semTools*. Retrieved from http://cran.r-project.org/web/packages/semTools/index.html

Karle, J. W., Watter, S., & Shedden, J. M. (2010). Task switching in video game players: Benefits of selective attention but not resistance to proactive interference. *Acta Psychologica, 134,* 70–78. http://dx.doi.org/10.1016/j.actpsy.2009.12.007

Kenny, D. A. (2015). *Measuring model fit*. Retrieved from http://davidakenny.net/cm/fit.htm

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological Bulletin, 136,* 849–874. http://dx.doi.org/10.1037/a0019842

Kleinsorge, T. (2004). Hierarchical switching with two types of judgment and two stimulus dimensions. *Experimental Psychology, 51,* 145–149. http://dx.doi.org/10.1027/1618-3169.51.2.145

Kleinsorge, T., & Heuer, H. (1999). Hierarchical switching in a multidimensional task space. *Psychological Research, 62,* 300–312. http://dx.doi.org/10.1007/s004260050060

Koch, I. (2001). Automatic and intentional activation of task sets. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1474–1486. http://dx.doi.org/10.1037/0278-7393.27.6.1474

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging, 15,* 126–147. http://dx.doi.org/10.1037/0882-7974.15.1.126

Kunde, W. (2001). Response-effect compatibility in manual choice reaction tasks. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 387–394. http://dx.doi.org/10.1037/0096-1523.27.2.387

Lawo, V., Philipp, A. M., Schuch, S., & Koch, I. (2012). The role of task preparation and task inhibition in age-related task-switching deficits. *Psychology and Aging, 27,* 1130–1137. http://dx.doi.org/10.1037/a0027455

Lee, K., Bull, R., & Ho, R. M. H. (2013). Developmental changes in executive functioning. *Child Development, 84,* 1933–1953. http://dx.doi.org/10.1111/cdev.12096

Lee, K., Ng, S. F., Pe, M. L., Ang, S. Y., Hasshim, M. N. A. M., & Bull, R. (2012). The cognitive underpinnings of emerging mathematical skills: Executive functioning, patterns, numeracy, and arithmetic. *British Journal of Educational Psychology, 82,* 82–99. http://dx.doi.org/10.1111/j.2044-8279.2010.02016.x

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21,* 59–80. http://dx.doi.org/10.1348/026151003321164627

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology, 49,* 764–766. http://dx.doi.org/10.1016/j.jesp.2013.03.013

Lien, M.-C., Schweickert, R., & Proctor, R. W. (2003). Task switching and response correspondence in the psychological refractory period paradigm. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 692–712. http://dx.doi.org/10.1037/0096-1523.29.3.692

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology, 6,* 1171. http://dx.doi.org/10.3389/fpsyg.2015.01171

Logan, G. D. (1985). Executive control of thought and action. *Acta Psychologica, 60,* 193–210. http://dx.doi.org/10.1016/0001-6918(85)90055-1

Logan, G. D., & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure? *Journal of Experimental Psychology: Human Perception and Performance, 29,* 575–599. http://dx.doi.org/10.1037/0096-1523.29.3.575

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review, 108,* 393–434. http://dx.doi.org/10.1037/0033-295X.108.2.393

Logan, G. D., Schneider, D. W., & Bundesen, C. (2007). Still clever after all these years: Searching for the homunculus in explicitly cued task switching. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 978–994. http://dx.doi.org/10.1037/0096-1523.33.4.978

Mayr, U. (2001). Age differences in the selection of mental sets: The role of inhibition, stimulus ambiguity, and response-set overlap. *Psychology and Aging, 16,* 96–109. http://dx.doi.org/10.1037/0882-7974.16.1.96

Mayr, U. (2002). Inhibition of action rules. *Psychonomic Bulletin & Review, 9,* 93–99. http://dx.doi.org/10.3758/BF03196261

Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General, 129,* 4–26. http://dx.doi.org/10.1037/0096-3445.129.1.4

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1124–1140. http://dx.doi.org/10.1037/0278-7393.26.5.1124

McCabe, D. P., Roediger, H. L., III, 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,* 222–243. http://dx.doi.org/10.1037/a0017619

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1423–1442. http://dx.doi.org/10.1037/0278-7393.22.6.1423

Meiran, N. (2000a). Modeling cognitive control in task-switching. *Psychological Research, 63,* 234–249. http://dx.doi.org/10.1007/s004269900004

Meiran, N. (2000b). Reconfiguration of stimulus task sets and response task sets during task switching. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 377–399). Cambridge, MA: MIT Press.

Meiran, N., Kessler, Y., & Adi-Japha, E. (2008). Control by action representation and input selection (CARIS): A theoretical framework for task switching. *Psychological Research, 72,* 473–500. http://dx.doi.org/10.1007/s00426-008-0136-8

Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development, 13,* 395–423. http://dx.doi.org/10.1080/15248372.2011.585478

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21,* 8–14. http://dx.doi.org/10.1177/0963721411429458

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,* 49–100. http://dx.doi.org/10.1006/cogp.1999.0734

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling, 9,* 599–620. http://dx.doi.org/10.1207/S15328007SEM0904_8

Nieuwenhuis, S., & Monsell, S. (2002). Residual costs in task switching: Testing the failure-to-engage hypothesis. *Psychonomic Bulletin & Review, 9,* 86–92. http://dx.doi.org/10.3758/BF03196259

Oberauer, K. (2005). Control of the contents of working memory—A comparison of two paradigms and two age groups. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 714–728. http://dx.doi.org/10.1037/0278-7393.31.4.714

Oberauer, K. (2009). Design for a working memory. In B. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 51). New York, NY: Academic Press. http://dx.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,* 1017–1045. http://dx.doi.org/10.1016/S0191-8869(99)00251-2

Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittmann, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence, 31,* 167–193. http://dx.doi.org/10.1016/S0160-2896(02)00115-0

Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittmann, W. W. (2008). Which working memory functions predict intelligence? *Intelligence, 36,* 641–652. http://dx.doi.org/10.1016/j.intell.2008.01.007

Oriet, C., & Jolcœur, P. (2003). Absence of perceptual processing during reconfiguration of task set. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 1036–1049. http://dx.doi.org/10.1037/0096-1523.29.5.1036

Pashler, H. (2000). Task switching and multitask performance. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 277–307). Cambridge, MA: MIT Press.

Philipp, A. M., & Koch, I. (2005). Switching of response modalities. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 58,* 1325–1338. http://dx.doi.org/10.1080/02724980443000656

Pornprasertmanit, S., Miller, P., Schoemann, A., Quick, C., & Jorgensen, T. D. (2016). *simsem: SIMulated structural equation modeling*. Retrieved from https://github.com/simsem/simsem

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114,* 510–532. http://dx.doi.org/10.1037/0033-2909.114.3.510

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Reimers, S., & Maylor, E. A. (2005). Task switching across the life span: Effects of age on general and specific switch costs. *Developmental Psychology, 41,* 661–671. http://dx.doi.org/10.1037/0012-1649.41.4.661

Rey-Mermet, A., & Meier, B. (2014). More conflict does not trigger more adjustment of cognitive control for subsequent events: A study of the bivalency effect. *Acta Psychologica, 145,* 111–117. http://dx.doi.org/10.1016/j.actpsy.2013.11.005

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 207–231. http://dx.doi.org/10.1037/0096-3445.124.2.207

Ropovik, I. (2014). Do executive functions predict the ability to learn problem-solving principles? *Intelligence, 44,* 64–74. http://dx.doi.org/10.1016/j.intell.2014.03.002

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2011). Modeling a cascade of effects: The role of speed and executive functioning in preterm/full-term differences in academic achievement. *Developmental Science, 14,* 1161–1175. http://dx.doi.org/10.1111/j.1467-7687.2011.01068.x

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48,* 1–36. http://dx.doi.org/10.18637/jss.v048.i02

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 763–797. http://dx.doi.org/10.1037/0096-1523.27.4.763

Rushworth, M. F. S., Passingham, R. E., & Nobre, A. C. (2002). Components of switching intentional set. *Journal of Cognitive Neuroscience, 14,* 1139–1150. http://dx.doi.org/10.1162/089892902760807159

Rushworth, M. F. S., Passingham, R. E., & Nobre, A. C. (2005). Components of attentional set-switching. *Experimental Psychology, 52,* 83–98. http://dx.doi.org/10.1027/1618-3169.52.2.83

Ruthruff, E., Remington, R. W., & Johnston, J. C. (2001). Switching between simple cognitive tasks: The interaction of top-down and bottom-up factors. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 1404–1419. http://dx.doi.org/10.1037/0096-1523.27.6.1404

Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology, 19,* 532–545. http://dx.doi.org/10.1037/0894-4105.19.4.532

Schneider, D. W., & Logan, G. D. (2006). Hierarchical control of cognitive processes: Switching tasks in sequences. *Journal of Experimental Psychology: General, 135,* 623–640. http://dx.doi.org/10.1037/0096-3445.135.4.623

Shaffer, L. H. (1965). Choice reaction with variable S-R mapping. *Journal of Experimental Psychology, 70,* 284–288. http://dx.doi.org/10.1037/h0022207

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 and Memory, 6,* 174–215. http://dx.doi.org/10.1037/0278-7393.6.2.174

Süß, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence, 30,* 261–288. http://dx.doi.org/10.1016/S0160-2896(01)00100-3

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., . . . Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51,* 247–250. http://dx.doi.org/10.1016/j.jml.2004.03.002

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence, 35,* 427–449. http://dx.doi.org/10.1016/j.intell.2006.09.001

Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *The British Journal of Educational Psychology, 82,* 100–119. http://dx.doi.org/10.1111/j.2044-8279.2011.02035.x

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin, 136,* 601–626. http://dx.doi.org/10.1037/a0019791

Vaughan, L., & Giovanello, K. (2010). Executive function in daily life: Age-related influences of executive processes on instrumental activities of daily living. *Psychology and Aging, 25,* 343–355. http://dx.doi.org/10.1037/a0017729

von Bastian, C. C., Locher, A., & Ruflin, M. (2013). Tatool: A Java-based open-source programming framework for psychological studies. *Behavior Research Methods, 45,* 108–115. http://dx.doi.org/10.3758/s13428-012-0224-y

von Bastian, C. C., & Oberauer, K. (2013). Distinct transfer effects of training different facets of working memory capacity. *Journal of Memory and Language, 69,* 36–58. http://dx.doi.org/10.1016/j.jml.2013.02.002

von Bastian, C. C., Souza, A. S., & Gade, M. (2016). No evidence for bilingual cognitive advantages: A test of four hypotheses. *Journal of Experimental Psychology: General, 145,* 246–258. http://dx.doi.org/10.1037/xge0000120

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology, 46,* 361–413. http://dx.doi.org/10.1016/S0010-0285(02)00520-0

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology, 4*(433), 1–22. http://dx.doi.org/10.3389/fpsyg.2013.00433

Willoughby, M. T., Blair, C. B., Wirth, R. J., & Greenberg, M. (2010). The measurement of executive function at age 3 years: Psychometric properties and criterion validity of a new battery of tasks. *Psychological Assessment, 22,* 306–317. http://dx.doi.org/10.1037/a0018708

Willoughby, M. T., Wirth, R. J., & Blair, C. B., & the Family Life Project Investigators. (2012). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment, 24,* 418–431. http://dx.doi.org/10.1037/a0025779

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 73,* 913–934. http://dx.doi.org/10.1177/0013164413495237

Appendix

Table A1
The Two Testing Orders Used in the Present Study

Test order 1		Test Order 2	
Shifting facet	Task	Shifting facet	Task
Block 1		*Block 1*	
Judgment	Two-Choice A	Dimension	Two-Choice B
Judgment	Four-choice	Dimension	Three-choice
Judgment	Three-choice	Dimension	Four-choice
Judgment	Two-Choice B	Dimension	Two-Choice A
Block 2		*Block 2*	
Response set	Two-Choice A	Mapping	Two-Choice B
Response set	Three-choice	Mapping	Three-choice
Response set	Four-choice	Mapping	Four-choice
Response set	Two-Choice B	Mapping	Two-Choice A
Block 3		*Block 3*	
Stimulus set	Two-Choice A	Stimulus set	Two-Choice B
Stimulus set	Three-choice	Stimulus set	Four-choice
Stimulus set	Four-choice	Stimulus set	Three-choice
Stimulus set	Two-Choice B	Stimulus set	Two-Choice A
Block 4		*Block 4*	
Mapping	Two-choice A	Response set	Two-Choice B
Mapping	Four-choice	Response set	Four-choice
Mapping	Three-choice	Response set	Three-choice
Mapping	Two-choice B	Response set	Two-Choice A
Block 5		*Block 5*	
Dimension	Two-Choice A	Judgment	Two-Choice B
Dimension	Four-choice	Judgment	Three-choice
Dimension	Three-choice	Judgment	Four-choice
Dimension	Two-Choice B	Judgment	Two-Choice A

Table A2
Descriptive Statistics for Error Rates

	Switch		Repetition	
Task	*M*	*SD*	*M*	*SD*
Judgment				
2A	.04	.05	.02	.03
2B	.03	.03	.03	.03
3	.05	.05	.04	.04
4	.04	.06	.03	.06
Dimension				
2A	.07	.06	.05	.04
2B	.05	.04	.03	.03
3	.09	.06	.08	.06
4	.05	.04	.03	.04
Mapping				
2A	.09	.07	.03	.03
2B	.10	.06	.04	.04
3	.07	.05	.04	.04
4	.09	.06	.04	.04
Response set				
2A	.08	.06	.03	.03
2B	.07	.06	.04	.04
3	.06	.05	.05	.04
4	.05	.05	.05	.05
Stimulus set				
2A	.03	.03	.03	.03
2B	.03	.03	.04	.03
3	.03	.03	.02	.02
4	.01	.01	.01	.02

Note. 2A = Two-Choice A; 2B = Two-Choice B; 3 = three-choice; 4 = four-choice.

Received July 12, 2016
Revision received April 21, 2017
Accepted April 25, 2017 ■