



The individual and situational factors predicting unethical behaviour in the workplace: a direct and conceptual replication of Jones & Kavanagh (1996)

Thomas Rhys Evans, James Edward Bartlett, Amelie Gourdon-Kanhukamwe, Tim Allen, Georgia Antoniou, Sara Arent, Dea Asmarani, Neelo Binns, Clara Boulton, Joshua Bowhay, Linnea Cederlund, Ramnik Chana, Yu Yang Chou, Kevina Dookheea, Gurbani Guglani, Abi Hamson-Ford, Holly Hood, Eci Balim Hostafci, Suha Jahanzeb, Kaylie Jones, Gemma Katie M Joseph, Tephilla Kemaless, Lamiss-Roya Karoui, Soha Khan, Renata Kviatkoyskyte, Jing Xuan Lee, Rahimah Lewis, Odelia Logan, Georgia Martin, Samaneh Vahedi Moghadam Masuleh, Rebecca McCulloch, Pragya Modi, Susan Nakro, Vy Nguyen, Minh Phan, Edward Philcox, Ana B Pliego, Sasha Raman, Anusha Vaithianathan Ramji, Anna M. H. Ramsay, Natasha E Rhone-Parkinson, Tayla Rizzo-Powell, Attila Simony, Roksana A Sobolak, Maia Thompson, Believe Ujobundo, Bhagyashree Vijayakumar, Natalja Wells-Dean, Michaela Natasha Wernick, Gemma Wood, Chan Young Yang, Zareen Zahid, Colleen Addicott, Hazel Bending, Elisabeth E. F. Bradford, Peter Edward Branney, Joanna Butler, Irene Chu, Oliver Clark, Laura Dean, Yannick Griep, Tom Hatton, Caroline Heard, Olivia Kowalczyk, Alexandra Lautarescu, Tony McCarthy, Kimberly R More, John Sabo, Joyce Elena Schleu, Dora Scholarios, Rosalind Searle, Ilham Sebah, Laura Wallis, Ruijie Wang, J  r  my E. Wilson-Lemoine, Emma Wilson-Lemoine & Calvin Burns

To cite this article: Thomas Rhys Evans, James Edward Bartlett, Amelie Gourdon-Kanhukamwe, Tim Allen, Georgia Antoniou, Sara Arent, Dea Asmarani, Neelo Binns, Clara Boulton, Joshua Bowhay, Linnea Cederlund, Ramnik Chana, Yu Yang Chou, Kevina Dookheea, Gurbani Guglani, Abi Hamson-Ford, Holly Hood, Eci Balim Hostafci, Suha Jahanzeb, Kaylie Jones, Gemma Katie M Joseph, Tephilla Kemaless, Lamiss-Roya Karoui, Soha Khan, Renata Kviatkoyskyte, Jing Xuan Lee, Rahimah Lewis, Odelia Logan, Georgia Martin, Samaneh Vahedi Moghadam Masuleh, Rebecca McCulloch, Pragya Modi, Susan Nakro, Vy Nguyen, Minh Phan, Edward Philcox, Ana B Pliego, Sasha Raman, Anusha Vaithianathan Ramji, Anna M. H. Ramsay, Natasha E Rhone-Parkinson, Tayla Rizzo-Powell, Attila Simony, Roksana A Sobolak, Maia Thompson, Believe Ujobundo, Bhagyashree Vijayakumar, Natalja Wells-Dean, Michaela Natasha Wernick, Gemma Wood, Chan Young Yang, Zareen Zahid, Colleen Addicott, Hazel Bending, Elisabeth E. F. Bradford, Peter Edward Branney, Joanna Butler, Irene Chu, Oliver Clark, Laura Dean, Yannick Griep, Tom Hatton, Caroline Heard, Olivia Kowalczyk, Alexandra Lautarescu, Tony McCarthy, Kimberly R More, John Sabo, Joyce Elena Schleu, Dora Scholarios, Rosalind Searle, Ilham Sebah, Laura Wallis, Ruijie Wang, J  r  my E. Wilson-Lemoine, Emma Wilson-Lemoine & Calvin Burns (27 Jun 2025): The individual and situational factors predicting unethical behaviour in the workplace: a direct and conceptual replication of Jones & Kavanagh (1996), *Comprehensive Results in Social Psychology*, DOI: [10.1080/23743603.2025.2507434](https://doi.org/10.1080/23743603.2025.2507434)

To link to this article: <https://doi.org/10.1080/23743603.2025.2507434>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 27 Jun 2025.



Submit your article to this journal [↗](#)



Article views: 254



View related articles [↗](#)



View Crossmark data [↗](#)

The individual and situational factors predicting unethical behaviour in the workplace: a direct and conceptual replication of Jones & Kavanagh (1996)

Thomas Rhys Evans^{a,b}, James Edward Bartlett^c, Amelie Gourdon-Kanhukamwe^d, Tim Allen^e, Georgia Antoniou^f, Sara Arent^g, Dea Asmarani^h, Neelo Binnsⁱ, Clara Boulton^g, Joshua Bowhay^e, Linnea Cederlund^j, Ramnik Chana^g, Yu Yang Chou^d, Kevina Dookheea^k, Gurbani Guglani^a, Abi Hamson-Ford^l, Holly Hood^a, Eci Balim Hostafci^a, Suha Jahanzeb^a, Kaylie Jones^m, Gemma Katie M Joseph^a, Tephilla Kemareshⁱ, Lamiss-Roya Karoui^c, Soha Khan^a, Renata Kviatkoyskyte^a, Jing Xuan Leeⁱ, Rahimah Lewis^a, Odelia Logan^a, Georgia Martin^a, Samaneh Vahedi Moghadam Masuleh^a, Rebecca McCullochⁱ, Pragya Modi^{l,n}, Susan Nakro^a, Vy Nguyenⁱ, Minh Phan^a, Edward Philcox^o, Ana B Pliego^a, Sasha Ramanⁱ, Anusha Vaithianathan Ramji^{p,q}, Anna M. H. Ramsay^r, Natasha E Rhone-Parkinson^a, Tayla Rizzo-Powell^a, Attila Simony^s, Roksana A Sobolak^a, Maia Thompson^a, Believe Ujobundo^t, Bhagyashree Vijayakumarⁱ, Natalja Wells-Dean^o, Michaela Natasha Wernick^h, Gemma Woodⁿ, Chan Young Yang^{u,aa}, Zareen Zahid^g, Colleen Addicott^k, Hazel Bending^e, Elisabeth E. F. Bradford^r, Peter Edward Branney^t, Joanna Butlerⁱ, Irene Chu^v, Oliver Clark^m, Laura Dean^h, Yannick Griep^{j†}, Tom Hattonⁿ, Caroline Heard^k, Olivia Kowalczyk^u, Alexandra Lautarescu^w, Tony McCarthyⁱ, Kimberly R More^{r,x}, John Sabo^f, Joyce Elena Schleu^j, Dora Scholariosⁱ, Rosalind Searle^c, Ilham Sebah^g, Laura Wallis^e, Ruijie Wang^l, Jérémy E. Wilson-Lemoine^{s,y}, Emma Wilson-Lemoine^z and Calvin Burns^{a,b}

^aSchool of Human Sciences, University of Greenwich, London, UK; ^bInstitute for Lifecourse Development, University of Greenwich, London, UK; ^cSchool of Psychology and Neuroscience, University of Glasgow, Glasgow, UK; ^dDepartment of Neuroimaging, King's College London, London, UK; ^eSchool of Health & Wellbeing, Plymouth Marjon University, Plymouth, UK; ^fSchool of Psychology and Life Sciences, Canterbury Christ Church University, Canterbury, UK; ^gDepartment of Psychology, Royal Holloway University of London, Egham, UK; ^hSheffield University Management School, The University of Sheffield, Sheffield, UK; ⁱDepartment of Work, Employment and Organisation, University of Strathclyde, Glasgow, UK; ^jBehavioural Science Institute, Radboud Universiteit, Nijmegen, Nijmegen, Gelderland, Netherlands; ^kSchool of Life and Medical Sciences, University of Hertfordshire, Hatfield, UK; ^lDepartment of Psychology, Bournemouth University, Dorset, Poole, UK; ^mDepartment of Psychology, Manchester Metropolitan University, Manchester, UK; ⁿSchool of Psychology, Arden University Ltd, Coventry, UK; ^oSchool of Psychology, Newcastle University, Newcastle upon Tyne, UK; ^pRutgers Cancer Institute, Rutgers University, New Brunswick, NJ, USA; ^qPsychosis Studies Department, King's College London, London, UK; ^rDepartment of Psychology, University of Dundee, Dundee, UK; ^sSchool of Psychology, University of East London, London, UK; ^tDepartment of Psychology, University of Bradford, West Yorkshire, Bradford, UK; ^uDepartment of Psychology, King's College London, London, UK; ^vNewcastle University Business School, Newcastle University, Newcastle upon Tyne, UK;

CONTACT Thomas Rhys Evans  thomas.evans@greenwich.ac.uk  School of Human Sciences, University of Greenwich, London, UK

[†]Present affiliation for Yannick Griep is Samergo, Rotterdam, the Netherlands and School of Industrial Psychology and Human Resource management, north-west university, Potchefstroom, South Africa.

Present affiliation for Jeremy Wilson-Lemoine is Department of Psychology, Texas State University, USA.

Present affiliation for Emma Wilson-Lemoine is Department of Psychology, University of Virginia, Charlottesville, VA, USA

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

^wDepartment of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; ^xDepartment for Health, University of Bath, Bath, UK; ^yESCP Business School, London, UK; ^zSchool of Mental Health and Psychological Sciences, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; ^{aa}Institute of Psychiatry, Psychology & Neuroscience, University College London, London, UK

ABSTRACT

Intentions to act unethically in the workplace are purported to be driven by a number of situational and individual factors. Across two seminal vignette experiments, Jones and Kavanagh reported inconsistent effect sizes for manager and peer influence and locus of control, consistent significant effects for work quality and Machiavellianism, and consistent non-significant effects for gender. Using an innovative multi-site collaboration, the current Registered Report represents a direct replication of these experiments ($N = 2218$), and adds a longitudinal conceptual replication capturing self-reported unethical work behaviour ($N = 1747$). Both replications found a consistent small effect of having a more external locus of control and male identity, and a consistent moderate effect of machiavellianism, for increasing unethical intentions and behaviour. The situational factors, whilst consistent in direction with that of the original study, varied more substantively in effect size. Our results highlight the value of multi-site collaborations and different replication types in developing conceptual, methodological, measurement and theoretical clarity to ensure future works can progress more rapidly to minimize the negative impacts of unethical workplace behaviour and improve individual's working lives. All materials, code and data for this project can be found here: osf.io/d3arx.

ARTICLE HISTORY

Received 15 March 2021

Accepted 5 May 2025

KEYWORDS

Unethical behaviour; workplace ethics; workplace influence; Machiavellianism; registered report

Introduction

From Governments to global businesses, unethical behaviour continues to represent a substantive social issue. Such is the extent that the 16th United Nations Sustainable Development Goal focuses upon building accountable institutions and tackling corruption (United Nations, 2015). The academic study of ethical decision-making continues to expand and contribute towards these goals, however relatively little is yet conclusive in how to adapt recruitment or training systems for individuals, or change an organisations' systems, structures or cultures, to minimise unethical workplace behaviour (Tenbrunsel & Smith-Crowe, 2008). Undermining our ability to make such developments is a consistent and robust body of evidence where the language is shared, findings are convergent, and evidence-quality is high. Inspired by such prevalent issues within the field, the current study reports both a direct and conceptual replication of Jones and Kavanagh (1996), studying the effects of individual and situational factors on unethical behaviour. Establishing the replicability of this cornerstone work represents an opportunity to re-evaluate the quality of evidence available and to establish new standards which may contribute to a more rigorous evidence base suitable for application.

Unethical work behaviour

Unethical work behaviours are defined as the “scalable actions and behaviours that employees engage in that deceive or exploit other persons or provide oneself (or one’s organization or associates) with an unfair advantage in the service of some other end” (Wiernik & Ones, 2018, p. 37). There are a wide range of workplace behaviours which have been labelled as unethical, including that of lying, cheating and stealing (Treviño et al., 2006). Unethical behaviour can occur through or beyond conscious reasoning and can be enacted to benefit the self or others (Veetkazhi et al., 2020).

Unethical workplace behaviour is often contextualized within models of Counterproductive Work Behaviour, defined as the “scalable actions and behaviours employees engage in that detract from [legitimate] organizational goals or well-being and include behaviours that bring about undesirable consequences for the organization or its stakeholders” (Ones & Dilchert, 2013, p. 645). Indeed, the wider counterproductive work behaviour label is popularly adopted as interchangeable with unethical behaviour by academics and practitioners alike, for example, in the rapid evidence assessment of the evidence on unethical workplace behaviour produced by the CIPD (2019). Whilst it is possible for a behaviour to be correctly defined as both unethical and counterproductive (Fida et al., 2015), the definitions refer to different groups of diverse behaviours with only small overlap and so should not be used interchangeably (Wiernik & Ones, 2018). The problematic conflation of definitions in this field has also occurred with the terms deviance and illegal behaviour.

This issue of semantics is problematic. As different sub-fields have adopted distinct terms and language, understanding of different dimensions of the field have been siloed (Kish-Gephart et al., 2010). This has immediate consequences for the interpretation and integration of such works, where it is ambiguous as to the extent to which different works/fields are referring to the same or different behaviours. As such, understanding of unethical workplace behaviour often suffers due to the jingle-jangle fallacy (Thorndike, 1904): the erroneous assumptions that the two behaviours are similar because they share the same label (jingle) or are different because they hold different labels (jangle). Incremental development of understanding is therefore obfuscated, and attempts to synthesise works can lead to diverse conclusions. For example, many conclusions of the Kish-Gephart et al. (2010) meta-analysis of unethical behaviour were significantly different to that of previous reviews which adopted wider counterproductive-based definitions. When attempting to draw meaningful consensuses for informing practice, such issues with definitions and differential conclusions can undermine influence.

A lack of definitional and conceptual clarity can also have a number of consequences for research through measurement practices and theory. For example, a diverse range of measurement scales and approaches have been developed to capture (un)ethical judgments, decision-making and behaviour (e.g. Pan & Sparks, 2012) and many existing measures do not clearly conform to the definitions or theories adopted (Wiernik & Ones, 2018). There have been a number of results which have been directly contradictory to the central theories adopted (Pan & Sparks, 2012). For example, counter to theories applied from social psychology where intentions drive behaviour (e.g. Theory of Planned Behaviour; Ajzen, 1985), and Rest’s (1986) awareness to judgment to intention to behaviour model of ethical decision-making, individual differences provide a consistently

stronger relationship to behaviour when compared to intentions (Kish-Gephart et al., 2010). As a consequence of the lack of coherence between definitions, theories, and measurement practices, “theories applied within this literature have guided our development of knowledge over time without careful integration, which is problematic for comparing results across contexts” (Mackey et al., 2021, p. 598).

The lack of clear conceptual/definitional focus within the field of unethical workplace behaviour, and subsequent inconsistencies in research practices, have had long-standing consequences for developing conceptual theory, establishing consensus in findings, and synthesizing findings to inform practice. However, whilst the specific findings tend to vary, there has been a consensus in the area of study of greatest importance to practice within the field. To-date, the majority of the evidence relating to unethical workplace behaviours has centred around identifying “bad apples” and “bad barrels” - the individual and contextual antecedents to unethical action (Kish-Gephart et al., 2010). Identifying and appreciating the extent of unique influence for each contributing factor is vital for informing change to minimize the prevalence of such problematic practices. However, there are many such models and studies considering different antecedents and dimensions of unethical practice and even within outcomes definitional/conceptual issues have led to inconsistent results. For example, in differentiating between unethical intentions and unethical actions (Rest, 1986).

In a review of literature between 1996 and 2003, O’Fallon and Butterfield (2005) identified 18 different individual factors explored within the extant literature, of which participant gender and value orientation were the most popularly studied. Similarly, 12 organisational factors were identified, with code of ethics and ethical climate/culture as the most popularly studied. Many of the factors listed continue to receive sporadic attention, particularly the organizational factors, and findings typically represent significant variation in effect size estimates (Bellé & Cantarelli, 2017). The origin of much of this diverse body of work is that of Jones and Kavanagh (1996), a highly influential paper featuring two vignette-based studies, early in the development of this field.

Jones and Kavanagh (1996)

Jones and Kavanagh (1996) conducted two near-identical experiments to explore the role of situational factors (quality of work experience, manager influence, peer influence) and individual factors (locus of control, Machiavellianism and participant gender) upon unethical behavioural intentions. Experiment 1 featured 138 upper-level undergraduate management students, and Experiment 2 included 154 MBA students. After capturing the individual factors through questionnaires, Jones and Kavanagh provided their students with a vignette about an organization where there was scope for unethical behaviour (to inflate expense claims). Participants were randomly allocated to different vignettes where the quality of work (high versus low) and behaviours of peers and a manager (ethical or unethical), were manipulated (see specific wording below in Table 2). Unethical behavioural intentions were then captured by asking participants to indicate their intentions to inflate their expense reports in this situation. Participants reported their intentions before completing manipulation check items asking about their hypothetical workplace, and demographic questions.

Table 1. Main effect sizes reported for experiment 1 and 2 (Jones & Kavanagh, 1996).

	Experiment 1 Beta Values (p-value)*	Experiment 2 Beta Values (p-value)	Experiment 1 Correlation Coefficient	Experiment 2 Correlation Coefficient
Quality of Work	.236 (<.001)	.153 (<.05)	.235 (<.01)	.149 (<.05)
Manager Influence	−1.371 (>.05)	.246 (<.001)	.066 (>.05)	.242 (<.001)
Peer Influence	−.007 (>.05)	.147 (<.05)	.247 (<.01)	.120 (>.05)
Locus of Control	.152 (<.05)	−.028 (>.05)	.211 (<.01)	.070 (>.05)
Machiavellianism	.181 (<.05)	.265 (<.001)	.216 (<.01)	.239 (<.01)
Gender	.049 (>.05)	.097 (>.05)	.033 (>.05)	.122 (>.05)

*These values have been distorted from the original through the introduction of a manager *peer interaction, particularly with respect to peer influence which was originally statistically significant ($p < .05$).

The results were concluded to be “relatively robust” (p. 520) across samples. Table 1 demonstrates the unstandardized Beta values and correlation coefficients for each of the six predictor variables studied in reference to the unethical behavioural intentions outcome. The correlational results between Experiment 1 and Experiment 2 should be mostly comparable given the use of identical methods and measurement practices with only minor differences in sample (Experiment 2 had a slightly older sample with more work experience) and wording of manipulation checks (which were captured after the data which is presented below).

Across the analyses, quality of work experience and Machiavellianism were consistently associated with behavioural intention, whilst small non-significant gender differences were reported. The strength of these relationships was relatively similar across the two experiments and provide mostly consistent evidence to clarify their role in the study of unethical behaviour. In contrast, manager influence, peer influence and locus of control report seemingly different strengths of effect (difference in $r > .1$ between experiments). For example, manager influence was meaningfully larger, and peer influence and locus of control was smaller, when comparing Experiment 2 to Experiment 1 results. As such, the two studies represent somewhat mixed results, although broad support was claimed for the notion that individual and situational variables are both important determinants of unethical behavioural intention.

Existing understanding

The body of work presented by Jones and Kavanagh (1996) began the proliferation of studies with a disparate range of predictor variables and diverse conceptualisations of unethical behaviour. Whilst the general role of both individual and contextual factors as antecedents of unethical behaviour has been accepted, O’Fallon and Butterfield’s (2005) review identified many inconsistencies in conclusions drawn across studies about the same antecedents.

For example, when considering gender as perhaps one of the most commonly measured individual factors, recent meta-analyses have concluded weak or small links, ranging from $r = .07$ (Mackey et al., 2021) to $r = .17$ (Pan & Sparks, 2012). Whilst initially encouraging, such meta-analytic estimates have obfuscated nuances within the results. For example, O’Fallon and Butterfield (2005) identified 23 studies with few or no differences, and 16 studies where women were significantly more ethical than men. Furthermore, individual conclusions have often differed, sometimes concluding gender

to “add nothing to understanding” (Kish-Gephart et al., 2010, p. 20) whereas others have reported meaningful differences in unethical behaviour (e.g. Gerlach et al., 2019; Pusch & Holtfreter, 2021). There is some evidence to suggest that controlling for social desirability diminishes this effect (Yang et al., 2017); however, no consensus has yet been established and as such a confident estimate of the effect is absent.

Similar false consensus has been established for Locus of Control, which represents the continuum of an individual's beliefs that outcomes are due to ability and effort (internal) or external factors like luck (external). Meta-analytic estimates range from $r = .06$ (Pan & Sparks, 2012) to $r = .13$ (Kish-Gephart et al., 2010) yet findings often reflect a range of non-significant effects, and positive correlations between external locus and unethical outcomes (O'Fallon & Butterfield, 2005). Similarly, meta-analyses of Machiavellianism, a trait associated with manipulating others for personal gain (Christie & Geis, 1970), have reported associations from $r = -.13$ (Pan & Sparks, 2012) to $r = -.27$ (Kish-Gephart et al., 2010), but there is evidence to suggest that the extent to which such traits are activated is relative to certain conditions, e.g. abusive supervision (Greenbaum et al., 2017), and thus highly variable.

Capturing the role of work quality and manager/peer influence has been slightly more complicated by the diverse number of terms which have been adopted to represent such factors. As such, the conclusions reached have been even more diverse. For example, Peng and Kim (2020) conclude ethical culture, represented by perceptions of rewards for ethical behaviour and senior management's role modelling of ethical behaviour, to reflect “the strongest predictor of counterproductive behavior” (p. 348), further supported by a meta-analysis in which it demonstrated incremental validity over a number of other leadership dimensions (T. W. H. Ng & Feldman, 2015). However, Pan and Sparks (2012) report the ethical environment, developed through manager and peer influence, to be weak and non-significant ($r = .08$). This becomes even more complex when considering recent research which suggests that such factors may interact, and that ethical employee behaviour may facilitate unethical behaviours from leaders (Ahmad et al., 2020). Finally, quality of work environment has been ill-defined and has also manifested in diverse conclusions. The most common conclusion suggests that job satisfaction is negatively related to unethical behaviour ($r = .24$, Kish-Gephart et al., 2010). The many ways in which work features can contribute to unethical behaviour have not yet been well identified and synthesised.

In sum, the very small and/or non-significant correlations reported by Jones and Kavanagh (1996) are typical of the literature where generally weak or null effects have been common across meta-analyses of all the themes captured. Whilst the meta-analyses reviewed appear to present coherent results, consistent with those produced by Jones & Kavanagh, we should not be fooled into thinking these have established reliable consensus from rigorous bodies of evidence (Mackey et al., 2021). Whilst not comprehensive of all the work completed in the field, the literature reviewed represents a range of conflicting conclusions to such an extent that it has been deemed “plagued by inconsistent findings” (Kish-Gephart et al., 2010, p. 17). There are a number of inconsistencies within papers (e.g. Jones & Kavanagh, 1996), across papers (e.g. Kish-Gephart et al., 2010; O'Fallon & Butterfield, 2005), and contradictions with theories (Pan & Sparks, 2012), and these have been of direct consequences for the integration of understanding. Indeed, the

field has been deemed to have produced “an inadequate understanding of the precise magnitude of relationships within workplace deviance’s nomological network” (Mackey et al., 2021, p. 598).

Attempts to resolve inconsistencies in conclusions for many antecedents of unethical behaviour are limited, with the broad body of evidence presenting a clear lack of consensus and significant heterogeneity in effect size estimates (Bellé & Cantarelli, 2017; O’Fallon & Butterfield, 2005; Tenbrunsel & Smith-Crowe, 2008). The wider field also has issues with establishing consensus due to conceptual ambiguity and measurement (Wiernik & Ones, 2018), where “the enormous body of workplace deviance research has developed without the cohesion necessary for nuanced insight” (Mackey et al., 2021, p. 598). Much of the conclusions drawn have been based on dated literature with low rigor (Treviño et al., 2006) and low sample sizes (Bellé & Cantarelli, 2017). The combination of inconsistent terms and conclusions, diverse methodological practices, and lack of replication work is especially problematic and dramatically limits the cumulative progress possible from research in the field.

The need for replication

Whilst replications exhaust important resources (Isager et al., 2023), replications of the Jones and Kavanagh (1996) studies would be relatively brief, require an easily accessible population, and secure substantive attention and impact for the field. Based upon the wide issues surrounding consistencies in definitions, findings and evidence-quality highlighted, and the nuances of this specific research, there are many reasons why replications of the Jones and Kavanagh (1996) experiments would be of great value.

First, the original manuscript postulated (but did not hypothesise) that individual and situational factors would interact to determine unethical behaviour. Whilst this is not an unreasonable expectation (Treviño et al., 2006), the study a) found little evidence of such an interaction effect beyond a single peer influence-manager influence interaction in Experiment 1, and b) the conclusions of simple main effects were problematic and inconsistent, likely due to low statistical power. The sample size of the two experiments reported were low ($n = 138/154$). The analyses had 80% power to detect correlations greater than $r = .24$ and $.22$, respectively, however across the two experiments only 3 of the 12 relationships between ethical intentions and the predictor variables of interest exceeded such a level, ranging from $r = .033$ to $.247$. Low power is an issue throughout the field, particularly for evaluations of antecedents from a combination of individual, team, and organizational levels simultaneously (Kish-Gephart et al., 2010; O’Boyle et al., 2011). High-powered replications would therefore be of great value to determine more specific and robust estimates of effect sizes.

Second, this work has been highly influential for the development of the field. Jones and Kavanagh (1996) have been cited over 500 times before 2020, were more recently reprinted in a book of the most influential research in business ethics (Michalos & Poff, 2012), and have featured in a number of reviews and meta-analyses on the ethical decision-making literature (e.g. Treviño et al., 2006). This popularity goes well beyond the 100 citations Makel et al. (2012) proposed as an expected threshold for replication attention. Furthermore, the ratio of citations to sample size, and many similar derivations of Isager’s et al. (2023) evaluation of replication priority through uncertainty of original

results and potential replication value, suggests that replication of the Jones and Kavanagh (1996) work should be high priority. The substantive number of citations is reflective of the central role this paper has had upon the development of the field, inspiring consideration of a wide range of situational and individual factors to predict unethical workplace behaviour.

Third, there have been no published replication works of the central project and relatively minimal studies of the same effects of interest. Whilst there have been a number of projects exploring various individual and situational factors to unethical decision-making, they infrequently provide converging conclusions. There are inconsistencies in findings between the two experiments reported by Jones and Kavanagh (1996), and with those reported in the extant literature. This is likely due to a lack of consistent definitions, criteria, and subsequent measurement practices, and replication attention. The results therefore represent a limited set of conclusions which lesser any potential impact. Here, a replication would be a high-value development to determine much more confident effect size estimates before driving developments through further conceptual replications to explore other factors, interactions between them, and draw wider claims of the generalizability of findings for practice.

Replications can be situated on a continuum based upon the extent to which the effect is examined in a similar way, ranging from exact to very far (LeBel et al., 2017). Hüffmeier et al. (2016) highlights the need for closer replications where there are low levels of subsequent evaluations and inconsistencies in results, whereby more conceptual replications have greater value for understanding generalizability and relevance for real-world application. The low levels of definitional clarity, replication efforts, and consistency in findings, yet high impact upon the development upon the field, would suggest that a close replication would be of great value to provide more robust estimates of the effects proposed.

There are also a number of factors surrounding this particular research which suggest that a conceptual replication would also be fruitful. For example, it might be reasonable to expect that answering questions on locus of control, Machiavellianism, and social desirability immediately before being presented with an ethically dubious scenario may have influenced participant's responding to the questions on unethical behavioural intentions (i.e. question order bias). Similarly, the use of hypothetical vignettes has been heavily criticised in favour of methods which may provide evidence more applicable for practice e.g. capturing real-world behavioural outcomes (Lonati et al., 2018). There are also concerns about the specific vignette adopted, in that the differences between conditions sometimes vary beyond the central factor of interest. For example, the prospected punishment is different between the unethical (fraud being noticed by payroll) and ethical (requiring more verification of expenses) peer influence conditions.

Furthermore, individual differences provide a consistently stronger relationship to behaviour than intentions (Kish-Gephart et al., 2010). This is possibly because unethical behaviours are a result of more impulsive decision-making rather than calculated responses (Kish-Gephart et al., 2010). The widespread use of vignettes and scenarios (Weber, 1992; Weber & Gillespie, 1998), although useful, may in some cases inadvertently prompt deliberation that takes participants out of the realm of these more realistic and impulsive types of decision-making. A conceptual replication adopting a self-report measure of real-world unethical workplace

behaviours could provide valuable insight into the extent to which different (and suboptimal) measurement approaches to unethical behaviour may be contributing to the inconsistencies in findings reported across the field.

The current study

Upon evaluation, the field provides little corroboration and mostly sporadic evidence, indicative of the need for greater replication attention. As such, a high-powered very close and conceptual replication of Jones and Kavanagh (1996) are proposed. The proposed replication studies are sufficiently powered to provide a confident estimate of each effect with high power and relatively low resource cost, and there is great potential impact from building a more confident evidence base. Such works aim to resolve inconsistencies in the presence and size of effects and provide a clearer basis for the study of main and interaction effects in the individual and situational predictors of intended and enacted unethical workplace behaviour.

Method (study 1 & 2)

All materials, ethics documentation, mock and (anonymised) raw data, meta-data and summary data, processing code, and analysis code is published on the Open Science Framework (osf.io/d3arx).

Recruitment of data-collection teams

Data was collected through 37 research groups. Recruitment of such groups was initiated in two ways. First, all Course Directors of Occupational and Applied Psychology courses running in the UK were emailed to encourage staff-student dyads to sign-up to collect data as part of their dissertation. Second, to maximise the likelihood that data collection targets were met, the lead author also made use of collaborative networks through social media and existing frameworks (e.g. Twitter) to identify data collection groups for Study 2.

Through all communications, interested parties were directed to the main OSF page for the project, which provided an overview of the project and timeframes, and details of how to sign up. Each data collection group had to include at least one student of any level (including undergraduate, postgraduate or doctorate), collect data in English, and to obtain ethical approval to conduct the study from their IRB unless their institution did not require approval, or the work could be covered by pre-existing approval. Based upon their resources, access to populations, and research interest, research groups could volunteer to collect data for either or both studies. All data collection teams had their contributions explicitly recognised through authorship of the final Stage 2 manuscript and details of each group can be found on the OSF project page.

Sample size

The original study featured two experiments with sample sizes of 138 and 154. A power analysis based upon the smallest effect size of interest (regression beta value) reported

was not feasible due to the incomprehensive reporting of unstandardized statistics. Instead, a power analysis was conducted based upon the smallest correlation observed between the unethical behavioural intention outcome and the three predictor variables with the greatest inconsistencies (manager influence, peer influence and locus of control, as reported above). The smallest correlation was with manager influence in Experiment 1 and was reported as $r = .066$. This was considered an acceptable smallest effect size of interest given that it also fell below all meta-analytic estimates of the relationships of interest as reviewed above (e.g. Kish-Gephart et al., 2010; Mackey et al., 2021).

Using a power calculation in R, as reported in the power analysis code document on the OSF project page, the number of participants required to observe this effect with 80% power and alpha of .05 is 1799. To account for the expected ~10% of careless responses (Meade & Craig, 2012), the current work aimed to collect data from a minimum of 2000 participants per study.

Whilst the use of multiple data collection groups can minimise issues relating to researcher degrees of freedom with respect to data stopping, the following stopping rules were applied. First, each data collection dyad were asked to collect completed responses from a minimum of 100, but no more than 250, participants (pre-exclusion for inattentive responding). Each data collection group was sent a weekly or fortnightly update email noting participant numbers pre- and post-exclusion. Second, data collection for each study was to completely end either on the 1st of August 2023, or when each group committed to data collection exceeded their minimum targets (more than 100 completions) with the total number of valid responses exceeding 2500), whichever came first. Data from all participants who completed the study, regardless of data-collection group, were analysed and made publicly available on the OSF.

Materials

The studies only ran in English and all data was collected online to maximise the likelihood that participants would more accurately report their unethical intentions or behaviours (Gnambs & Kaspar, 2015). Qualtrics links were set up for data collection groups based upon a central template to minimize the likelihood of inconsistencies. Only the ethical documentation before and after the study content differed between data collection sites.

Data management

For each study, data from all data-collection groups were collated and analysed as a single dataset to minimize type 1 errors.

Analyses mostly replicated the same structure provided in the original studies and were conducted through R and R Studio as prepared in the preregistered analysis scripts. The packages *apaTables* (Stanley & Spence, 2018), *dplyr* (Wickham et al., 2020), *ggcorrplot* (Kassambara, 2019), *Hmisc* (Harrell, 2021), *lme4* (Bates et al., 2015), *plyr* (Wickham, 2011), *psych* (Revelle, 2021), *pwr* (Champely, 2020) and *userfriendlyscience* (Peters, 2018) were used.

Criteria for replication

Each of the effects associated with the six independent variables are considered replicated in the context of the LeBel et al. (2019) replication criteria. The results were considered “consistent” where confidence intervals overlapped with the main effects previously reported in each of the original two experiments (for the direct replication) or the direct replication (for the conceptual replication). Similarly, where the confidence intervals did not overlap with 0, a signal was considered “detected.”

Study 1 – Direct Replication

Procedure & design

An online experimental design was adopted. Participants were asked to complete the experimental vignette study of Jones and Kavanagh (1996) as originally designed. After providing consent, participants first completed questionnaires capturing locus of control, Machiavellianism and social desirability. Participants were then randomly allocated to one of eight manipulated vignettes and asked to report their intention to inflate an expense request. These vignettes purportedly differed in three ways: whether they reported a low or high-quality work experience, ethical or unethical peer influence, and ethical or unethical manager influence (see Table 2). Participants were then asked to complete manipulation check questions and note their demographic details before being debriefed.

Participants were randomly allocated into one of the eight different possible combinations of conditions using the randomizer function in Qualtrics. Evenly presented elements were adopted to prevent unequally weighted groups. As demographic questions were presented at the end of the study to prevent influence upon the data, it was not possible to allocate proportional numbers of all gender identities to each group.

Materials

Locus of Control, Machiavellianism, and Social Desirability – Concurrent with the original study, ten items of the Rotter (1966), Christie and Geis (1970) and Crowne and Marlowe (1960) scales (respectively) were adopted using a 5-point likert scale response. The 10 items adopted from each of the original scales were provided upon request by personal communication with one of the original authors. Internal reliability estimates, as assessed by Cronbach’s Alpha, were originally reported in Experiment 1 as .76, .72 and .56, respectively, and were unreported for Experiment 2.

The vignettes used and the various conditions are reported in Table 2.

Unethical Behavioural Intention was captured by four items assessed through a 5-point likert scale. Again, the specific wording of items was specified by personal communication with the original author upon request. Internal reliability was originally estimated through Cronbach’s Alpha ($\alpha = .87$).

Table 2. Vignettes and study conditions.

<p>Irving Brockbank International is a fast-paced, reputable advertising firm in the Northeast. You have been with the firm as an accounts executive for about six months. Among other responsibilities, you consult with smaller advertising agencies promoting local area businesses. This requires that you periodically visit your clients, and you typically travel about once a month.</p>	
<p>High Quality Work Experience: Since working for the firm, you have been given a nice office, a competitive salary and good benefits. The company values you as an employee, and your immediate supervisor has been very appreciative of your efforts. In general, you have been very satisfied with the quality of your employment in the firm.</p>	<p>Low Quality Work Experience: Since working for the firm, you've discovered that your salary is below industry average, the benefits are not very good, and your "office" is a tiny cubicle you share with 2 other people. The company doesn't seem to value you as an employee, and your immediate supervisor has not been very appreciative of your efforts. In general, you have not been very satisfied with the quality of your employment in the firm.</p>
<p>After coming back from a trip, you show your expense report to one of your friends in the office who basically does the same job you do. After looking at your expense report, he:</p>	
<p>Ethical Peer Influence: raises his eyebrows, and says "you need to be very careful in reporting your expenses. For instance, I see you spent £30.00 for the cab fare from the airport. That could look very suspicious to payroll. I know that some people in the company do inflate their trip report expenses, but most people are pretty accurate. If we aren't, payroll will start requiring even more verification of expenses. Plus, company travel funds will get depleted, and it'll be harder to justify a trip."</p>	<p>Unethical Peer Influence: laughs and says, "You actually put down what you really spent!" He then says to you, "Let me tell you, most of us add a few pounds here and there – its sort of a private perk." He then proceeds to show you which expenses you could increase (for example taxi fares, meals and tips) and why (for example, a receipt is not required). He also tells you by how much you could increase expenses, without it being noticed by payroll as excessive.</p>
<p>Ethical Managerial Influence: Later, you happen to see the expense report of your boss from when he did the same trip about a month ago. You notice that his expenses are reported as much less than you reported. For example, you went ahead and put down estimates of meals of about £40.00 per day. Your boss reported that he spent about £15.00 per day. Obviously, the boss is very meticulous and accurate in reporting trip expenses.</p>	<p>Unethical Managerial Influence: Later, you happen to see the expense report of your boss from when he did the same trip about a month ago. You notice his expenses are reported as much greater than you know it really cost, from having just taken the same trip. For example, you stayed in the same hotel as he did and it cost you £7.00 cab fare between the hotel and airport. He reported that it cost £20.00 each way. You notice other expenses as well that are most likely inflated.</p>

Three manipulation check items were included and assessed through three response options each, in accordance with the original experiment and as specified by personal communication with the original author.

A range of demographic questions were included to describe the sample and support secondary analyses and generalizability statements. Gender, age, ethnicity, employment status, months of full-time work experience, industry in which they had the majority of their work experience, academic field of study, and year of study was captured.

Participants

Data was collected through 26 research groups across the UK and represented a total sample size expected to be between 2000 and 2500. The original study used students in their a) final year of an undergraduate management course (experiment 1) and b) MBA evening students (experiment 2). To balance the demands of sample size with the need for a comparable population, and as research with students tend to report slightly larger effect sizes on research in ethical judgements to that of the general public (Pan & Sparks, 2012), the current study recruited only

UK-based management- and psychology-based students from any year of their undergraduate or postgraduate study. Recruitment was predominantly through university research participation schemes (e.g. SONA) and targeted messages through virtual learning environments (e.g. Moodle). Specific details on the methods each data collection group adopted are available on the meta-data documents on the OSF project page.

Exclusion

Our study was attempted by 4477 participants. First, we considered participants who did not finish the survey ($N = 275$) as withdrawn. All questions forced a response to ensure no missing data. We then removed all data for which the study meta-data flagged pilot or spam data by Qualtrics ($N = 1168$). The final sample pre-exclusions therefore included complete data from 3034 participants.

In the original studies, manipulation checks were analysed but did not form the basis of exclusion. As such, to capture inattentive engagement, in line with good practices (i.e. Meade & Craig, 2012), participants were excluded if they respond “No – do not use my data” to the following: “We want to ensure that all data we analyse is meaningful and accurately reflects engagement with the study content. In your honest opinion, should we use your data in our analyses of this study? Your response to this will have no impact or consequences to you so please do be honest.” We excluded 816 participants on this basis, resulting in a final sample of 2218.

Analysis plan

First, three Pearson’s Chi-squared tests were conducted to explore the efficacy of the respective experimental manipulations on quality of work experience, peer influence and manager influence. Second, descriptive statistics (means and standard deviations) and omega coefficients were presented within a correlation matrix.

Third, to test the primary hypotheses, a multiple regression was conducted. The categorical independent variables (quality of work experience, peer influence, managerial influence and gender) were dummy coded before being regressed upon behavioural intention alongside Machiavellianism and locus of control total scores. The original manuscript inconsistently reported a manager-peer influence interaction so to minimize the likelihood of type 1 errors and prevent postulation of effects that were not originally hypothesized, interaction effects were not studied.

Finally, in addition to replicating the same set of analyses provided for the original experiments, analyses were conducted to acknowledge the multi-site nature of data collection. First, the current study visualised the consistency of relationship strength across data collection groups by creating a heatmap-inspired correlation table. Second, a linear mixed-effects regression model was conducted, using data collection group number as a random effect and quality of work experience, peer influence, managerial influence, gender, Machiavellianism, and locus of control as fixed effects.

Deviations from preregistration

- In the Stage 1 Registered Report we did not explicitly note that data collected during the period of piloting by the data collection teams, or indeed spam completions, would be excluded. We applied this additional exclusion criterion to the direct replication because each data collection team piloted their own survey links multiple times (which was considered good to role-model best practice) and due to practical concerns surrounding data quality (e.g. some data collection teams recruited participants through social media and raised concerns, verified by the meta-data and Qualtrics system, that bots were completing the survey to secure the financial compensation).
- The configuration of data collection groups sometimes deviated from the agreed protocol due to practicality. As such, dyads were sometimes triads or tetrads, and the minimum sample required from each group was not always met.

Study 2 – Conceptual Replication

Procedure & design

A longitudinal online questionnaire design was adopted, adopting self-reported unethical workplace behaviour as the outcome. At time 1, participants completed demographic questions, followed by a battery of questionnaires on specific psychological constructs representing the themes originally studied. At time 2, between 7 and 28 days following initial participation, participants declared their use of unethical workplace behaviours from the last 12 months. At both time points, participants were asked to confirm whether their data was meaningful using the same engagement question as used in Study 1 (Meade & Craig, 2012).

Materials

A range of demographic questions were included to describe the sample and support secondary analyses and generalizability statements. Gender, age, ethnicity, student status, months of full-time work experience, and industry in which they had the majority of their work experience were captured.

Locus of Control was captured through the Work Locus of Control Scale (Spector, 1988) which has 16 items on a 6-point likert and was deemed most suitable as a widely adopted and context-specific measure of locus of control (T. W. Ng et al., 2006).

Machiavellianism was presented through the Mach IV (Christie & Geis, 1970). There were twenty items responded to on a 7-point likert, some of which were altered from the original to refer to “people” rather than “men” (Miller et al., 2019). Due to the concerns surrounding factor structure, the full 20-item scale was adopted and used to create a total score, allowing the opportunity for subsequent secondary analyses to consider alternative structures (e.g. Rauthmann, 2013).

Social Desirability was measured through the Balanced Inventory of Desirable Responding Short Form (Hart et al., 2015). This 16-item scale differentiates between self-

deceptive enhancement and impression management, and using a 7-point likert produces data with good psychometric properties, unlike many short forms of alternative measures (Barger, 2002).

Quality of Work Experience was considered through the lens of job satisfaction. This was captured through the Job in General scale (Ironson et al., 1989), chosen due the applicability to any job type and preferable psychometric properties of the data it produces compared to alternative short job satisfaction measures (Van Saane et al., 2003). Eighteen adjectives to describe a participants work e.g. pleasant and enjoyable, were rated and scored according to original instructions: No (0),? (1) or Yes (3).

Manager Influence was assessed using the 10-item Ethical Leadership Questionnaire (Brown et al., 2005). Participants rated their manager on a 7-point likert ranging from “Strongly Disagree” to “Strongly Agree.” This scale is the most widely adopted assessment of this construct, with other measures often considered to be dominated by items which fall outside the scope of ethical leadership specifically (Wiernik & Ones, 2018).

Peer Influence was captured through five questions exploring participants’ observations or knowledge of peers’ unethical workplace behaviours. The questions represent the Unethical Behaviour Toward Employees subscale developed by Kaptein (2008), chosen due to its applicability to a wide variety of workplaces and industries, and clear differentiation in focus from the outcome measurement. Participants responded on a 7-point likert ranging from “Never” to “(Almost) always.”

Unethical Workplace Behaviours were captured at time 2 through participants’ declarations of use of 17 unethical behaviours as developed by Newstrom and Ruch (1975) and commonly attributed to Akaah (1992). These items were deemed the most relevant due to their applicability to remote and face-to-face work, and a critical review of the field which declared the items to be “the most conceptually clear and comprehensive measure of employee unethical behaviours” (Wiernik & Ones, 2018, p. 42). Responses were captured through a 7-point likert scale with only anchors of “Never” and “Frequently,” in-line with practices which maximise self-reporting of unethical work behaviours (Wouters et al., 2014).

Participants

Data was collected through 17 research groups and represented a total sample size expected to be between 2000 and 2500. To facilitate wider claims of generalizability, the current study used an opportunity sampling technique to recruit UK and US participants who were employed at the time of data collection. Recruitment was through a variety of strategies including the use of university research participation schemes (e.g. SONA) and targeted messages through virtual learning environments (e.g. Moodle), social media, and collaborative research networks. Specific details on the methods each data collection group adopted are available on the meta-data documents on the OSF project page.

Exclusion

First, we considered participants who did not finish the first survey as withdrawn. All questions forced a response to ensure no missing data. Following the check for inattentive engagement, 2791 participants were invited to complete the second survey. Of these, 1044 did not complete or were withdrawn for inattentive engagement, leaving a final sample of 1747.

Analysis plan

First, descriptive statistics (means and standard deviations) and omega coefficients were presented alongside a correlation matrix. Second, as before, a multiple regression was conducted to predict unethical workplace behaviours from the time 1 measures. To further acknowledge the role of differences between data collection sites, a heatmap-inspired correlation table, and a linear mixed effects regression model with a “data collection group” random effect, were again reported.

Deviations from preregistration

- The two deviations reported for the direct replication were also followed for this conceptual replication.
- Based upon difficulties with data collection, a 1-month extension to data collection and stage 2 submission was agreed with the Editor. The editor agreed to use of Prolific.ac to collect a significant portion of data for the conceptual replication, instead of extending the project timeline a further year. As such, data collection continued until the end of September 2023 rather than August 2023. Data was slightly below the minimum expected sample size (1747/1799) but this was considered acceptable given the size of deviation and agreed timeframes of the project.
- One team (25) had an unintentional error whereby their quality of work experience assessment allowed participants to pick more than one response per question. To manage this, as participants could score either No (0),? (1) or Yes (3), all conflicting scores were recoded as “1” to acknowledge a mixed response. This influenced the scores of 10 participants in total.

Analyses

Direct replication

All analyses represent the preregistered analysis plans as originally intended, with a final sample of 2218. This included 1454 who identified as female, 711 individuals who identified as male, one who identified as non-binary, 42 who preferred not to identify and 10 who did not provide gender information. Ages ranged from 18 to 71 ($M = 25.6$, $SD = 8.19$). A large proportion of the sample identified as white ($N = 1348$) or Asian/Asian British ($N = 544$). However, our sample also included individuals identifying as Black, African, Caribbean or Black British ($N = 152$), Mixed/Multiple Ethnicities ($N = 101$), and Other Ethnic Groups ($N = 63$). Of those who reported their student status, the sample was predominantly first year ($N = 650$), only/final year ($N = 530$), or second year ($N = 449$).

Table 3. Direct replication: means, standard deviations, and correlations (* = $p < .05$; ** = $p < .01$).

Variable	<i>M</i>	<i>SD</i>	<i>Omega</i>	1	2	3	4	5	6	7
1. Quality of Work Environment	0.49	0.50								
2. Manager Influence	0.50	0.50		-.01						
3. Peer Influence	0.51	0.50		.01	-.01					
4. Locus of Control	28.42	4.90	.72	-.01	.01	.04				
5. Machiavellianism	26.20	4.93	.74	.01	.01	.03	.33**			
6. Gender	0.67	0.47		.00	-.01	-.02	-.01	-.12**		
7. Social Desirability	30.51	4.31	.50	-.02	-.06*	.01	-.25**	-.39**	-.01	
8. Unethical Behavioural Intentions	9.93	3.86	.90	-.05*	.07**	.06*	.14**	.29**	-.15**	-.20**

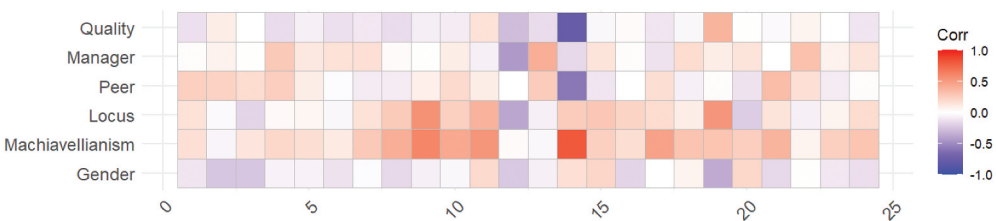


Figure 1. Heatmap inspired visualization of correlations to unethical behavioural intentions across data collection groups.

students. Of the 2192 who responded, 1215 were currently employed, and this group held an average of 35 months of work experience ($SD = 90$).

First, three Pearson’s Chi-squared tests were conducted to determine whether participants allocated to the different manipulated vignette conditions responded differently to the manipulation check items. For quality of work experience ($\chi^2 (2, N = 2218) = 586.84, p < .001$), peer influence ($\chi^2 (2, N = 2218) = 283.18, p < .001$) and manager influence ($\chi^2 (2, N = 2218) = 742.89, p < .001$), significant differences in responses between groups in the associated manipulation checks are congruent with, and provide initial evidence for, a successful manipulation of these factors. [Appendix A](#) shows bar plots for each manipulation check question against its respective condition.

Table 4. Direct replication regression results using unethical behavioural intentions as the criterion.

Predictor	<i>b</i>		<i>beta</i>	<i>beta</i>		<i>sr</i> ²		<i>r</i>	Fit
	<i>b</i>	95% CI [LL, UL]		95% CI [LL, UL]	<i>sr</i> ²	95% CI [LL, UL]			
(Intercept)	3.87**	[2.74, 4.99]							
Quality of Work Experience	−0.40*	[−0.71, −0.09]	−0.05	[−0.09, −0.01]	.00	[−.00, .01]	−.05*		
Manager Influence	0.53**	[0.23, 0.84]	0.07	[0.03, 0.11]	.00	[−.00, .01]	.07**		
Peer Influence	0.34*	[0.04, 0.65]	0.04	[0.00, 0.08]	.00	[−.00, .01]	.06*		
Locus of Control	0.04*	[0.01, 0.07]	0.05	[0.01, 0.09]	.00	[−.00, .01]	.14**		
Machiavellianism	0.20**	[0.17, 0.24]	0.26	[0.22, 0.30]	.06	[.04, .08]	.29**		
Gender	−1.00**	[−1.33, −0.67]	−0.12	[−0.16, −0.08]	.01	[.01, .02]	−.15**		
<i>R</i> ² = .112**									
95% CI [.0913]									

Note. A significant *b*-weight indicates the beta-weight and semi-partial correlation are also significant. *b* represents unstandardized regression weights. *beta* indicates the standardized regression weights. *sr*² represents the semi-partial correlation squared. *r* represents the zero-order correlation. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively.

* indicates $p < .05$. ** indicates $p < .01$.

Second, descriptive statistics are presented (Table 3) within a correlation matrix. Please note, from this point, participants who did not report a male or female identity were removed. Statistically significant ($p < .05$) relationships were found between the unethical behavioural intentions outcome and all predictors. The extent to which the strength of relationships vary across data collection groups can be found in Figure 1. Furthermore, internal reliability estimates were highly congruent with that of the original study, including a low value for social desirability.

Third, a multiple regression was conducted and reported in full below in Table 4. Signals were detected for all variables as no confidence intervals overlapped with 0, however the proportion of variance in unethical intentions explained was small (11%). Findings were generally mixed when compared to the original study. Machiavellianism (as expected) and gender (unexpectedly) were the two strongest predictors, whilst the contribution of variance explained by quality of work experience and locus of control was much more modest than originally reported. A linear mixed-effects regression model, to recognise potential inconsistencies in results across data collection groups, had very small variance (1.07) and standard deviation (1.03) estimates around the random intercept of data collection team, and the model provided largely unaltered estimates except for gender which was weaker (see Appendix B).

Conceptual replication

All analyses represent the preregistered analysis plans as originally intended, with a final sample of 1747. This includes 1001 females, 738 males, and eight individuals who identified with neither label. The sample was predominantly from the UK ($N = 1692$), non-students ($N = 1493$) with a mean age of 36.9 ($SD = 11.7$). Of the UK sample, most identified as White ($N = 1303$), although the sample also included individuals identifying as Asian/Asian British ($N = 146$), Black, African, Caribbean or Black British ($N = 118$), Mixed/Multiple Ethnicities ($N = 73$), and Other Ethnic Groups ($N = 52$). For the US-based participants ($N = 55$), 12 identified as Hispanic, Latino or of Spanish origin. US participants were most commonly identified as White ($N = 33$), but the sample also included individuals identifying with Other groups ($N = 8$), Asian ($N = 4$), two or more races ($N = 4$), Black or African American ($N = 2$) and American Indian or Alaska Native ($N = 1$). All participants were employed, with a mean 172 months of work experience ($SD = 255$). Data from the 8 individuals not identifying with a male or female identity were removed before analyses.

Descriptive statistics for all variables are presented below (Table 5) within a correlation matrix. As to be expected from the sample size, statistically significant ($p < .05$)

Table 5. Conceptual replication: means, standard deviations, and correlations (* = $p < .05$; ** = $p < .01$).

Variable	<i>M</i>	<i>SD</i>	<i>Omega</i>	1	2	3	4	5	6	7
1. Quality of Work Experience	38.98	12.75	.94							
2. Manager Influence	49.07	12.49	.95	.54**						
3. Peer Influence	9.21	5.89	.88	-.31**	-.40**					
4. Locus of Control	48.41	10.41	.89	-.38**	-.32**	.17**				
5. Machiavellianism	75.54	11.29	.81	-.31**	-.30**	.23**	.46**			
6. Gender	0.58	0.49		.05*	.01	.02	-.05*	-.14**		
7. Social Desirability	62.49	12.63	.85	.17**	.17**	-.13**	-.24**	-.44**	-.06*	
8. Unethical Workplace Behaviour	31.68	12.37	.92	-.14**	-.19**	.30**	.13**	.30**	-.15**	-.32**

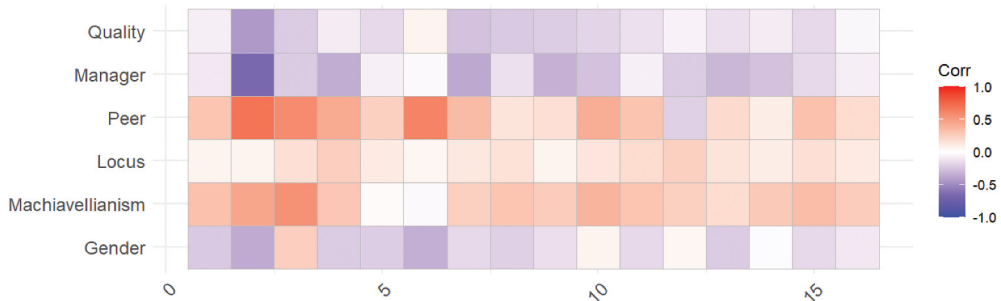


Figure 2. Heatmap inspired visualization of correlations to unethical behaviour across data collection groups (excluding one team who collected 2 participants).

Table 6. Conceptual replication: regression results using unethical behaviour as the criterion.

Predictor	<i>b</i>	<i>b</i> 95% CI [LL, UL]	<i>beta</i>	<i>beta</i> 95% CI [LL, UL]	<i>sr</i> ²	<i>sr</i> ² 95% CI [LL, UL]	<i>r</i>	Fit
(Intercept)	11.50**	[5.72, 17.28]						
Quality of Work Experience	0.02	[−0.03, 0.08]	0.03	[−0.03, 0.08]	.00	[−.00, .00]	−.14**	
Manager Influence	−0.04	[−0.09, 0.02]	−0.04	[−0.09, 0.02]	.00	[−.00, .00]	−.19**	
Peer Influence	0.51**	[0.41, 0.61]	0.24	[0.20, 0.29]	.05	[.03, .07]	.30**	
Locus of Control	−0.03	[−0.09, 0.03]	−0.03	[−0.08, 0.02]	.00	[−.00, .00]	.13**	
Machiavellianism	0.26**	[0.20, 0.32]	0.24	[0.19, 0.29]	.04	[.02, .06]	.30**	
Gender	−3.07**	[−4.17, −1.98]	−0.12	[−0.17, −0.08]	.01	[.00, .03]	−.15**	
<i>R</i> ² = .163**								
95% CI [.1319]								

Note. A significant *b*-weight indicates the beta-weight and semi-partial correlation are also significant. *b* represents unstandardized regression weights. *beta* indicates the standardized regression weights. *sr*² represents the semi-partial correlation squared. *r* represents the zero-order correlation. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively. * indicates $p < .05$. ** indicates $p < .01$.

relationships were found between the unethical behaviour outcome and all predictors. The extent to which the strength of relationships varies across data collection groups can be found in Figure 2. Furthermore, internal reliability estimates were generally high and similar to previous reports of such properties.

Third, a multiple regression was conducted and reported in full below in Table 6. A signal was not detected for quality of work experience, manager influence and locus of control, as confidence intervals overlapped with 0, however the proportion of variance in unethical intentions explained was slightly larger than the direct replication (16%). Again, mixed results were observed. Peer influence and gender (unexpectedly) and machiavellianism (as expected) predicted modest proportions of variance in unethical behaviour. A linear mixed-effects regression model, to recognise potential inconsistencies in results across data collection groups, had small variance (2.38) and standard deviation (1.54) estimates around the random intercept of data collection team, providing very minor and insignificant deviations from these estimates (see Appendix B).

Discussion

Across two high-power replications, this study explored the role of individual and situational factors on unethical behavioral intentions in the workplace. Our findings represent a mixed picture of effects, as might have been expected from the lack of consensus in the extant literature (Mackey et al., 2021). Across the replications we find a consistent small effect of having a more external locus of control and male identity, and a consistent moderate effect of machiavellianism, for increasing unethical intentions and behaviour. The situational factors, whilst consistent in direction with that of the original study, varied more substantively in effect size. As interpreted below, our results point to the need for more attention on conceptual, methodological, measurement and theoretical clarity to ensure future works can progress more directly to achieve the purpose of research in this domain – to minimize the negative impacts of unethical workplace behaviour and improve individual's working lives.

Individual effects

Our study found that lower quality of work experience was associated with greater unethical behaviour in the workplace. The size of this effect was especially modest when observed in the context of unethical intentions. Quality of work was defined as Job Satisfaction by Jones and Kavanagh (1996) and in both our current replications. There are differences in components of satisfaction covered, however. For example, the original vignette adopted contains dimensions of work environment (nice office/tiny cubicle), salary (competitive/below industry average), company/supervisor appreciation and general perceptions. The conceptual replication covered a more general affective evaluation of satisfaction, and in other studies, quality of work experience has been conceptualised completely differently, e.g. as an organizational environment. Askew et al. (2015) state that organizational environment is the most important antecedent of unethical behavior, though in their study it seems to reflect the “ethical climate, culture, and codes of conduct established within an organization.” These factors seem to all be important antecedents to unethical behaviour as outlined by Kish-Gephart et al. (2010), but job satisfaction was characterised as an individual characteristic in their study (rather than a situational characteristic), while organizational environment qualities fell under organizational environment characteristics. Differentiating between objective workplace characteristics and subsequent perceptions of agreeableness will offer greater insights into how the individual and the environment interact to drive unethical intentions.

Whilst not the strongest predictor of unethicality, as was originally argued by Peng and Kim (2020), our results provide further evidence for the influence of leadership and seniors' role-modeling upon employee behaviour (Michell et al., 2022; T. W. H. Ng & Feldman, 2015). Unethicality of management was a significant predictor of unethical intention and behaviour, although most strongly for the latter. These results suggest that unethical behaviour may represent more impulsive decision-making (Kish-Gephart et al., 2010) and/or employees may more readily copy the actions of senior colleagues (Kaptein, 2011), but perhaps the underlying cognitive processes (i.e. intent) may be less readily manipulated by others' actions. Combined with other small effects (e.g. Pan &

Sparks, 2012) and recent interactions (e.g. Ahmad et al., 2020), a more structured and detailed evaluation of the different levels of power and influence across individuals, teams and structures likely represents a fruitful avenue of future exploration.

The significant main effect of peer influence is an example of social influence in organisational settings, and is in line with Izraeli (1988) who suggested that employees who witness unethical behaviour by their colleagues are more likely to engage in unethical behaviour themselves. As with manager influence, the influence of peers was more modest in influencing intentions than behaviours. Likewise, the original study by Jones and Kavanagh (1996) found a significant effect of peer influence in Experiment 1 but the effect was then “masked by the significant interaction between peer and managerial influence” (p. 520) in Experiment 2. As such, these results provide further encouragement of consideration of more social factors of the workplace and thus may also benefit from a more structured analysis of influence.

As identified in earlier studies, including Jones and Kavanagh (1996), an external locus of control is associated with unethicity – believing outcomes are out of their control leads people to engage in unethical behavior as they do not feel as much responsibility or control for their actions. The Jones and Kavanagh (1996) study was inconsistent in its findings with regard to individual variables, as only experiment one found a main effect of locus of control. However, here, we found main effects of almost identical size in both direct and conceptual replications. Previous research by Trevino (1986) and Valentine et al. (2019) suggest that situational influences (e.g. the presence of an ethics code) can have an impact in reducing the externality of an individuals’ locus of control. Greater recognition of this interaction between situational and personal factors is a priority. Furthermore, the debate around whether locus of control is a unique construct or part of a larger construct related to self-evaluation (Galvin et al., 2018) remains contentious (Judge & Bono, 2001; Judge et al., 2002). Clear contextualisation of work locus of control should be considered, as associated variables like self-efficacy are likely to influence unethical behaviors in combination with locus of control and subsequently influence the size or direction of effects (Valentine et al., 2019).

In line with previous research (Collins & Schmidt, 1993; Jones & Kavanagh, 1996), our replications are consistent in arguing that machiavellianism represents a strong predictor of unethical workplace behaviour. It is unsurprising that if individuals are more likely to engage in manipulative and self-serving behaviours across social and interpersonal contexts, they are also more likely to engage in unethical workplace behaviour. This work provokes an important question as to whether machiavellianism is relevant for all the different kinds of unethical behaviour that can be observed and we encourage future researchers to reflect upon this beyond pro-self unethical behaviours. Furthermore, these findings should be interpreted in context of the strength of correlations ($r = .39$ and $.44$) observed between machiavellianism and social desirability, which may be indicative of participant bias and/or self-deceptive enhancement.

Finally, the results of the two replications were highly convergent in suggesting a very modest role of gender, similar in size of effect to that of previous reviews (Pan & Sparks, 2012). Concurrent with much of the extant literature that identifies differences, individuals identifying as males report slightly greater unethical workplace behaviours and intentions compared to those identifying as women. Estimates for individuals who wished to identify

Table 7. Comparison of correlations across all studies (some directions have been reversed for consistency/clarity).

	Present Study		Jones and Kavanagh (1996)		LeBel et al. (2019) Criteria	
	Direct Replication (N = 2218)	Conceptual Replication (N = 1747)	Experiment 1 (N = 138)	Experiment 2 (N = 154)	Signal Detected?	Consistent?
Quality of Work	−.05 (−.09, −.01)	−.14 (−.18, −.09)	−.24	−.15	Yes	Conceptual with Exp 2 only.
Manager Unethicality	.07 (.03, .11)	.19 (.14, .23)	.07	.24	Yes	Direct with Exp 1 only.
Peer Unethicality	.06 (.01, .10)	.30 (.26, .34)	.25	.12	Yes	No.
External Locus of Control	.14 (.10, .18)	.13 (.09, .18)	.21	.07	Yes	No.
Machiavellianism	.29 (.25, .33)	.30 (.26, .35)	.22	.24	Yes	No.
Gender (Male=1, Female=0)	.15 (.11, .19)	.15 (.10, .20)	.03	.12	Yes	Both replications with Exp 2 only.

with neither label were much less certain due to the small sample sizes and were therefore not analysed due to the risk of drawing misleading and imprecise conclusions. It would be of significant value to examine effect sizes in larger samples of individuals identifying with other gender identities, and to explore which dimensions of gender drive this effect.

Evidence synthesis

There are a number of important observations when examining the profile of the results as a whole (see Table 7). For example, across our replications, we found converging evidence on the strength of the relationship between unethical intentions and locus of control, machiavellianism and gender. This convergence is high (~.01 difference in correlation coefficient) and on one hand might be considered unsurprising given the similarity in measurement strategies of these variables between the two replication types (i.e. similar or identical questionnaires). On the other hand, this result might still be considered surprising by the differences in outcomes measured: the direct replication captured hypothetical intentions to inflate expense reports and the conceptual replication captured a range of real-world unethical behaviours.

There are meaningful differences in estimates of effect size between the direct and conceptual replication where measurement strategies changed from manipulated vignette to questionnaire. For these three variables (quality of work, manager influence and peer influence), all the relationships were stronger for the questionnaire-based conceptual replication capturing real-world actions. This set of results may be indicative of the influence of measurement effects upon effect size, and looks to represent a fruitful avenue for both understanding previous inconsistencies and future priorities.

Based upon our results, individual factors seem to more consistently play a role in influencing unethical behavioral intentions specifically, but the proportion of variance

explained is low and main effects were small in size. This difference in effects sizes between individual and situational factors was not observed for the conceptual/questionnaire-based replication. As such, results may indicate that existing works using vignette-based methodologies or experimentally manipulated variables are likely to underestimate relationship sizes for real-world interpretation, however it is unclear whether this effect is limited to just contextual or situational factors.

Finally, given the closeness of methods and samples between the direct replication and the original experiments (Jones & Kavanagh, 1996), we expected the estimates of effects from the direct replication to be closer to the original than those established by the conceptual replication. For the situational factors where effect sizes did not closely converge between replication types, estimates from the conceptual replication were closer to the effect sizes originally reported than the direct replication. This could be indicative of a number of explanations, including changes in working conditions, practices and standards over the last two decades (e.g. introduction of the Bribery Act 2010 in the UK), inflated original estimates based upon small sample sizes, a lessened influence of manipulation in a sample of (mostly) psychology students who are experienced with such methods, or the greater similarity in type of sample (more work-experienced).

Project limitations and future research

Our replications captured both intentions (direct replication) and behaviours (conceptual replication), with a clear profile of similarities and differences between results. Concerns have long been raised surrounding the measurement of behavioural intentions being conflated with actual behaviour (Trevino & Victor, 1992). Studies consistently reveal a discrepancy between an individual's intentions and their behaviour, with this applying vice versa also (Lawler, 1970). As such, differentiating intentions from behaviour may provide crucial in interpreting the extant literature due to the inconsistencies in results and measurement approaches. Furthermore, there are other factors to consider such as the role of intention strength (which is characterised by features such as intention stability) in negotiating the unethical intention-unethical behavior gap. For example, if intention strength is poor despite the presence of unethical intentions, the unethical behavior might not be carried out. Conner and Norman (2022) suggest that “moderators of the intention-behavior relationship may be mediated through intention stability (and perhaps other features of intention strength).” Such concepts may help further unpick the relationship between unethical behavioural intentions and behaviours and to determine the extent to which the two distinct bodies of evidence on these outcomes can be consistently united.

There are increasingly nuanced discussions surrounding the replication crisis which highlight the role of theory, and often the lack of robust theory, as a core contributor to failed replications or inconsistencies across bodies of evidence (Oberauer & Lewandowsky, 2019). As such, we encourage future works to engage more with theoretical frameworks (for example, the Theory of Planned Behaviour (Ajzen, 1985)), and believe this is particularly valuable in the context of differentiating between intentions and behaviours. This may provide a more structured approach to elucidate as to why some of the effects differed between the two outcome conceptualisations, and why individual predictors may be more consistent than situational/contextual factors.

Although the conceptual replication aimed to capture real-world unethical behaviour and the measures were implemented to resolve some of the well-established concerns surrounding use of hypothetical vignettes, questionnaires are still problematic for representing real-life scenarios and thus further limit the generalisability of the findings. The current study may have used the most theoretically aligned tool (Wiernik & Ones, 2018), but these are over two decades old and are unlikely to represent the full range of unethical behaviours evidenced in modern society. This does, however, represent wider issues in the field surrounding the measurement of real-world unethical behaviors, and raises the question of what the most accurate and ecologically valid method to measure them would be. Furthermore, we captured no qualitative data to assess timing, pertinence, reasoning, or other behaviours and/or experiences. Qualitative data may be particularly critical in resolving inconsistencies by encouraging a richer contextual outlook on how perceptions and intentions for ethical behaviours in the workplace are formed and influenced.

Defining, identifying and measuring unethical workplace behaviours has been problematic in academic and real-world practice with inconsistencies in the use of measurements, tools, and theories (Wiernik & Ones, 2018). Reinterpreting the existing literature with a consistent perspective on the defining features of ethical behaviour in a workplace context would provide a clearer foundation for the study of their potential impacts upon employees. Currently, all studies, including the current work, fail to make clear statements about the extent to which their findings can apply only to a specific given behaviour considered unethical, or rather a collection of behaviours (Treviño et al., 2006). Little attention or focus has been placed upon identifying where and why different behaviours may have different antecedents and consequences, and a clearer position on these factors may contribute to a more progressive and impactful body of work capable of informing action.

Given the methods used to collect data, caution should be exercised when attempting to generalise the current findings. The sample size was large and recruiting participants from universities was advantageous, providing flexibility and access to a large pool of participants. However, for the direct replication, for example, participants were predominantly undergraduate students (*first year* ($N = 650$), *only/final year* ($N = 530$) and *second year* ($N = 449$)) who are often entry-level workers in the workplace. Indeed, only 55% of the sample were employed, and they had an average of >3 years work experience. Importantly, applying these hypothetical responses to a vignette should be a far distance from claims about how these effects may work in practice. Furthermore, females represented a larger proportion of the sample. For example, the total number of female participants ($N = 1454$) from the direct replication population was double the total of male participants ($N = 711$). Similar concerns could be raised about the representation of ethnicity groups and age. As the samples predominantly originated in the UK, it seems likely that the participants may be quite homogenous, and it raises the question of cross-cultural validity of our findings (Henrich et al., 2010). Indeed, the direct replication revealed that the majority who reported their ethnicity identified as white ($N = 1348$) or Asian/Asian British ($N = 544$). Indeed, the participants from the conceptual experiment were predominantly from the UK ($N = 1692$) and identified as White ($N = 1303$). As such, the several sample qualities discussed suggest generalisability may be somewhat limited.

The exploration of the antecedents and consequences of unethical workplace behaviour has been diverse and inconsistent. There are relatively few examples of

clear replications, and little consistent examination of the extant literature to determine whether inconsistencies are due to different theories, designs, definitions, or measurement approaches. This diverse evidence base is likely indicative of the wider emphasis in our research quality upon novelty. Given the extent of inconsistencies identified, we encourage researchers to engage in less aimless hypothesis testing (Scheel et al., 2020) and accept the need for more meta-work to establish clearer foundations and address bigger questions. Placing more value, and investing more time, in building coherent, theory-informed, bodies of evidence is likely to be a critical factor for increasing influence upon practice. Only after this will it be possible to study the complex interactions, for example, between individual factors and situational/contextual factors. To provide just one example, we might expect contract types to have an especially complex association with unethical behaviour, reflecting interactions between various personal factors like financial and job insecurity (Jiang et al., 2022), and environmental factors like opportunities for unethicality and career progression. The original Jones and Kavanagh (1996) work considered some post-hoc interactions, but the extant literature is lacking in a structured and consistent account of how the various individual and contextual factors may interact. We suggest that following a more descriptive mapping through meta-work, future research should prioritise recognising the complex interplay and interactions between individuals' cognitions, behavioural intention/attitudes/perceptions and the behaviours themselves.

Applications

Findings from this study predominantly emphasise the individual factors driving unethical workplace behaviour. Primarily, this reinforces the need for rigorous recruitment to positions (e.g. considering negative personality traits like machiavellianism when selecting individuals for influential or high-risk roles) and maximising person-job fit. However, acknowledging Trait Activation Theory (Tett & Burnett, 2003) whereby certain contexts may more easily cue certain traits, it is important not to dismiss the role of the context, or indeed the accountability of the employer in providing a healthy workplace. The results of the current research, in context of the existing literature, suggest that more incisive developments will be made only with a much more nuanced coverage of interactions between the individuals and structures.

By providing further insight into the factors that define and promote unethical behaviours in the workplace, this research aims to support employees and managers alike. For example, findings from this study could be directed towards use in amending employee well-being programmes or workplace support interventions to better address the underlying needs behind pro-self unethical behaviour. We hope that the valuable datasets produced can support further insights into the dynamics of unethical workplace behaviour, and that the quantity of data collected supports the estimation of more consistent and rigorous effect size estimates as the number of reviews and meta-analyses in this important domain continues to grow. Furthermore, in discussing the limitations of our project, and the broader field at length, we hope to give researchers reason to pause and reflect upon definitions, theories, measurement approaches, and other such factors in the

hope that it helps build a more rigorous and impactful body of evidence from which policy and practice can be informed.

On a broader level, this work also demonstrates the under-utilized potential of offering meaningful research opportunities for students, having engaged 37 student-led teams to collect high-quality data from nearly 4000 participants, predominantly over a single academic year. There are many structured programmes which support this type of learning (e.g. CREP; Wagge et al., 2019) and the current study evidences what is possible with a Stage 1 Registered Report and a modest local network. The open scholarship practices like preregistration evidenced throughout this project embody the critical skills most universities purport to promote (e.g. scientific literacy and confidence; Pownall, Azevedo, et al., 2023). Despite anecdotal evidence of this value from the students contributing to this project, there is limited evidence evaluating to what extent open practices like preregistration contribute to student outcomes (Pownall, Pennington, et al., 2023). As such, we encourage ambitious researchers to follow us in leading rich and meaningful research experiences for students, working together to harness and nurture their skills and passions. Furthermore, we also encourage researchers to go further in embedding evaluation of the impact upon their students' personal and professional development.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Institute for Lifecourse Development, University of Greenwich.

ORCID

Thomas Rhys Evans  <http://orcid.org/0000-0002-6670-0718>

Pragya Modi  <http://orcid.org/0009-0000-9696-7547>

References

- Ahmad, M. G., Klotz, A. C., & Bolino, M. C. (2020). Can good followers create unethical leaders? How follower citizenship leads to leader moral licensing and unethical behavior. *Journal of Applied Psychology*, 106(9), 1374–1390. <https://doi.org/10.1037/apl0000839>
- Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), *Action control: From cognition to behavior* (pp. 11–39). Springer-Verlag.
- Akaah, I. P. (1992). Social inclusion as a marketing ethics correlate. *Journal of Business Ethics*, 11(8), 599–608. <https://doi.org/10.1007/BF00872271>
- Askew, O. A., Beisler, J. M., & Keel, J. (2015). Current trends of unethical behavior within organizations. *International Journal of Management & Information Systems (IJMIS)*, 19(3), 107–114. <https://doi.org/10.19030/ijmis.v19i3.9374>
- Barger, S. D. (2002). The Marlowe-Crowne Affect Scale: Short forms, psychometric structure, and social desirability. *Journal of Personality Assessment*, 79(2), 286–305. https://doi.org/10.1207/S15327752JPA7902_11

- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bellé, N., & Cantarelli, P. (2017). What causes unethical behavior? A meta-analysis to set an agenda for public administration research. *Public Administration Review*, 77(3), 327–339. <https://doi.org/10.1111/puar.12714>
- Brown, M. E., Treviño, L. K., & Harrison, D. A. (2005). Ethical leadership: A social learning perspective for construct development and testing. *Organizational Behavior and Human Decision Processes*, 97(2), 117–134. <https://doi.org/10.1016/j.obhdp.2005.03.002>
- Champely, S. (2020). *Pwr: Basic functions for power analysis*. R package version 1.3-0. <https://CRAN.R-project.org/package=pwr>
- Christie, R., & Geis, F. L. (1970). *Studies in machiavellianism*. Academic.
- CIPD. (2019). *Rotten apples, bad barrels and sticky situations: A rapid evidence assessment of the research literature on unethical workplace behaviour*. https://www.cipd.co.uk/Images/unethical-behaviour-in-the-workplace_technical-report-1_tcm18-57020.pdf
- Collins, J. M., & Schmidt, F. L. (1993). Personality, integrity, and white collar crime: A construct validity study. *Personnel Psychology*, 46(2), 295–311. <https://doi.org/10.1111/j.1744-6570.1993.tb00875.x>
- Conner, M., & Norman, P. (2022). Understanding the intention-behavior gap: The role of intention strength. *Frontiers in Psychology*, 13, 923464. <https://doi.org/10.3389/fpsyg.2022.923464>
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24(4), 349–354. <https://doi.org/10.1037/h0047358>
- Fida, R., Paciello, M., Tramontano, C., Fontaine, R. G., Barbaranelli, C., & Farnese, M. L. (2015). An integrative approach to understanding counterproductive work behavior: The roles of stressors, negative emotions, and moral disengagement. *Journal of Business Ethics*, 130(1), 131–144. <https://doi.org/10.1007/s10551-014-2209-5>
- Galvin, B. M., Randel, A. E., Collins, B. J., & Johnson, R. E. (2018). Changing the focus of locus (of control): A targeted review of the locus of control literature and agenda for future research. *Journal of Organizational Behavior*, 39(7), 820–833. <https://doi.org/10.1002/job.2275>
- Gerlach, P., Teodorescu, K., & Hertwig, R. (2019). The truth about lies: A meta-analysis on dishonest behavior. *Psychological Bulletin*, 145(1), 1–44. <https://doi.org/10.1037/bul0000174>
- Gnams, T., & Kaspar, K. (2015). Disclosure of sensitive behaviors across self-administered survey modes: A meta-analysis. *Behavior Research Methods*, 47(4), 1237–1259. <https://doi.org/10.3758/s13428-014-0533-4>
- Greenbaum, R. L., Hill, A., Mawritz, M. B., & Quade, M. J. (2017). Employee machiavellianism to unethical behavior: The role of abusive supervision as a trait activator. *Journal of Management*, 43(2), 585–609. <https://doi.org/10.1177/0149206314535434>
- Harrell, F. E., Jr. (2021). *Hmisc: Harrell Miscellaneous*. R package version 4.5-0. <https://CRAN.R-project.org/package=Hmisc>
- Hart, C. M., Ritchie, T. D., Hepper, E. G., & Gebauer, J. E. (2015). The balanced inventory of desirable responding short form (BIDR-16). *Sage Open*, 5(4), 2158244015621113. <https://doi.org/10.1177/2158244015621113>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83. <https://doi.org/10.1017/S0140525X0999152X>
- Hüffmeier, J., Mazei, J., & Schultze, T. (2016). Reconceptualizing replication as a sequence of different studies: A replication typology. *Journal of Experimental Social Psychology*, 66, 81–92. <https://doi.org/10.1016/j.jesp.2015.09.009>
- Ironson, G. H., Smith, P. C., Brannick, M. T., Gibson, W. M., & Paul, K. B. (1989). Construction of a job in General scale: A comparison of global, composite, and specific measures. *Journal of Applied Psychology*, 74(2), 193–200. <https://doi.org/10.1037/0021-9010.74.2.193>
- Isager, P. M., Van Aert, R., Bahník, Š., Brandt, M. J., DeSoto, K. A., Giner-Sorolla, R., Krueger, J. I., Perugini, M., Ropovik, I., van 't Veer, A. E., Vranka, M., & Lakens, D. (2023). Deciding what to replicate: A decision model for replication study selection under resource and knowledge constraints. *Psychological Methods*, 28(2), 438–451. <https://doi.org/10.1037/met0000438>

- Izraeli, D. (1988). Ethical beliefs and behavior among managers: A cross-cultural perspective. *Journal of Business Ethics*, 7(4), 263–271. <https://doi.org/10.1007/BF00381831>
- Jiang, L., Lawrence, A., & Xu, X. (2022). Does a stick work? A meta-analytic examination of curvilinear relationships between job insecurity and employee workplace behaviors. *Journal of Organizational Behavior*, 43(8), 1410–1445. <https://doi.org/10.1002/job.2652>
- Jones, G. E., & Kavanagh, M. J. (1996). An experimental examination of the effects of individual and situational factors on unethical behavioral intentions in the workplace. *Journal of Business Ethics*, 15(5), 511–523. <https://doi.org/10.1007/BF00381927>
- Judge, T. A., & Bono, J. E. (2001). Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology*, 86(1), 80–92. <https://doi.org/10.1037/0021-9010.86.1.80>
- Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of Personality & Social Psychology*, 83(3), 693–710. <https://doi.org/10.1037/0022-3514.83.3.693>
- Kaptein, M. (2008). Developing a measure of unethical behavior in the workplace: A stakeholder perspective. *Journal of Management*, 34(5), 978–1008. <https://doi.org/10.1177/0149206308318614>
- Kaptein, M. (2011). Understanding unethical behavior by unraveling ethical culture. *Human Relations*, 64(6), 843–869. <https://doi.org/10.1177/0018726710390536>
- Kassambara, A. (2019). *Ggcorrplot: Visualization of a correlation matrix using 'ggplot2'*. R package version 0.1.3. <https://CRAN.R-project.org/package=ggcorrplot>
- Kish-Gephart, J. J., Harrison, D. A., & Treviño, L. K. (2010). Bad apples, bad cases, and bad barrels: Meta-analytic evidence about sources of unethical decisions at work. *Journal of Applied Psychology*, 95(1), 1–31. <https://doi.org/10.1037/a0017103>
- Lawler, E. E. (1970). Job attitudes and employee motivation: theory, research, and practice 1. *Personnel Psychology*, 23(2), 223–237. <https://doi.org/10.1111/j.1744-6570.1970.tb01652.x>
- LeBel, E. P., Berger, D., Campbell, L., & Loving, T. J. (2017). Falsifiability is not optional. *Journal of Personality & Social Psychology*, 113(2), 254–261. <https://doi.org/10.1037/pspi0000106>
- LeBel, E. P., Vanpaemel, W., Cheung, I., & Campbell, L. (2019). A brief guide to evaluate replications. *Meta-Psychology*, 3, 1–9. <https://doi.org/10.15626/MP.2018.843>
- Lonati, S., Quiroga, B. F., Zehnder, C., & Antonakis, J. (2018). On doing relevant and rigorous experiments: Review and recommendations. *Journal of Operations Management*, 64(1), 19–40. <https://doi.org/10.1016/j.jom.2018.10.003>
- Mackey, J. D., McAllister, C. P., Ellen, B. P., III, & Carson, J. E. (2021). A meta-analysis of interpersonal and organizational workplace deviance research. *Journal of Management*, 47(3), 597–622. <https://doi.org/10.1177/0149206319862612>
- Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science*, 7(6), 537–542. <https://doi.org/10.1177/1745691612460688>
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17(3), 437–455. <https://doi.org/10.1037/a0028085>
- Michalos, A. C., & Poff, D. C. (2012). *Citation classics from the journal of business ethics: Celebrating the first thirty years of publication (vol. 2)*. Springer Science & Business Media.
- Miller, B. K., Nicols, K., & Konopaske, R. (2019). Measurement invariance tests of revisions to archaically worded items in the mach IV scale. *PLOS ONE*, 14(10), e0223504. <https://doi.org/10.1371/journal.pone.0223504>
- Newstrom, J. W., & Ruch, W. A. (1975). *The ethics of management and the management of ethics* (Vol. 23). Michigan State University Business Topics (Winter).
- Ng, T. W., Sorensen, K. L., & Eby, L. T. (2006). Locus of control at work: A meta-analysis. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 27(8), 1057–1087. <https://doi.org/10.1002/job.416>

- Ng, T. W. H., & Feldman, D. C. (2015). Ethical leadership: Meta-analytic evidence of criterion-related and incremental validity. *Journal of Applied Psychology*, 100(3), 948–965. <https://doi.org/10.1037/a0038246>
- O'Boyle, E. H., Forsyth, D. R., & O'Boyle, A. S. (2011). Bad apples or bad barrels: An examination of group-and organizational-level effects in the study of counterproductive work behavior. *Group & Organization Management*, 36(1), 39–69. <https://doi.org/10.1177/1059601110390998>
- O'Fallon, M. J., & Butterfield, K. D. (2005). A review of the empirical ethical decision-making literature: 1996–2003. *Journal of Business Ethics*, 59(4), 375–413. <https://doi.org/10.1007/s10551-005-2929-7>
- Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic Bulletin & Review*, 26(5), 1596–1618. <https://doi.org/10.3758/s13423-019-01645-2>
- Ones, D. S., & Dilchert, S. (2013). Counterproductive work behaviors: Concepts, measurement, and nomological network. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I.-C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbook of testing and assessment in psychology, Vol. 1. test theory and testing and assessment in industrial and organizational psychology* (pp. 643–659). American Psychological Association.
- Pan, Y., & Sparks, J. R. (2012). Predictors, consequence, and measurement of ethical judgments: Review and meta-analysis. *Journal of Business Research*, 65(1), 84–91. <https://doi.org/10.1016/j.jbusres.2011.02.002>
- Peng, A. C., & Kim, D. (2020). A meta-analytic test of the differential pathways linking ethical leadership to normative conduct. *Journal of Organizational Behavior*, 41(4), 348–368. <https://doi.org/10.1002/job.2427>
- Peters, G. (2018). *Userfriendlyscience: Quantitative analysis made accessible*. R package version 0.7.2. <https://doi.org/10.17605/osf.io/txequ>
- Pownall, M., Azevedo, F., König, L. M., Slack, H. R., Evans, T., Flack, Z., Grinschgl, S., Elsherif, M. M., Gilligan-Lee, K. A., de Oliveira, C. M. F., Gjoneska, B., Kalandadze, T., Button, K., Ashcroft-Jones, S., Terry, J., Albayrak-Aydemir, N., Dëchtërenko, F., Alzahawi, S. . . . Sladekova, M. (2023). Teaching open and reproducible scholarship: A critical review of the evidence base for current pedagogical methods and their outcomes. *Royal Society Open Science*, 10(5), 221–255. <https://doi.org/10.1098/rsos.221255>
- Pownall, M., Pennington, C. R., Norris, E., Juanchich, M., Smailes, D., Russell, S., & Clark, K. (2023). Evaluating the pedagogical effectiveness of study preregistration in the undergraduate dissertation. *Advances in Methods and Practices in Psychological Science*, 6(4), 25152459231202724.
- Pusch, N., & Holtfreter, K. (2021). Individual and organizational predictors of white-collar crime: A meta-analysis. *Journal of White Collar and Corporate Crime*, 2(1), 5–23. <https://doi.org/10.1177/2631309X19901317>
- Rauthmann, J. F. (2013). Investigating the MACH-IV with item response theory and proposing the trimmed MACH. *Journal of Personality Assessment*, 95(4), 388–397. <https://doi.org/10.1080/00223891.2012.742905>
- Rest, J. (1986). *Development in judging moral issues*. University of Minnesota Press.
- Revelle, W. (2021). *Psych: Procedures for personality and psychological Research*. R package version = 2.1.6. <https://CRAN.R-project.org/package=psych>
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General & Applied*, 80(1), 1–28. <https://doi.org/10.1037/h0092976>
- Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2020). Why hypothesis testers should spend less time testing hypotheses. *Perspectives on Psychological Science*, 16(4), 744–755. <https://doi.org/10.1177/1745691620966795>
- Stanley, D. J., & Spence, J. R. (2018). Reproducible tables in psychology using the apaTables package. *Advances in Methods and Practices in Psychological Science*, 1(3), 415–431. <https://doi.org/10.1177/2515245918773743>
- Spector, P. E. (1988). Development of work locus of control scale. *Journal of Occupational Psychology*, 61(4), 335–340. <https://doi.org/10.1111/j.2044-8325.1988.tb00470.x>
- Tenbrunsel, A. E., & Smith-Crowe, K. (2008). 13 ethical decision making: Where we've been and where we're going. *Academy of Management Annals*, 2(1), 545–607. <https://doi.org/10.5465/19416520802211677>

- Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, 88(3), 500–517. <https://doi.org/10.1037/0021-9010.88.3.500>
- Thorndike, E. L. (1904). *An introduction to the theory of mental and social measurements*. The Science Press.
- Trevino, L. K. (1986). Ethical decision making in organizations: A person-situation interactionist model. *Academy of Management Review*, 11(3), 601–617. <https://doi.org/10.2307/258313>
- Trevino, L. K., & Victor, B. (1992). Peer reporting of unethical behavior: A social context perspective. *Academy of Management Journal*, 35(1), 38–64. <https://doi.org/10.2307/256472>
- Treviño, L. K., Weaver, G. R., & Reynolds, S. J. (2006). Behavioral ethics in organizations: A review. *Journal of Management*, 32(6), 951–990. <https://doi.org/10.1177/0149206306294258>
- United Nations. (2015). *Resolution adopted by the general assembly on 25 September 2015, transforming our world: The 2030 agenda for sustainable development*. <https://sdgs.un.org/2030agenda>
- Valentine, S. R., Hanson, S. K., & Fleischman, G. M. (2019). The presence of ethics codes and employees' internal locus of control, social Aversion/Malevolence, and ethical judgment of incivility: A study of smaller organizations. *Journal of Business Ethics*, 160(3), 657–674. <https://doi.org/10.1007/s10551-018-3880-8>
- Van Saane, N., Sluiter, J. K., Verbeek, J. H. A. M., & Frings-Dresen, M. H. W. (2003). Reliability and validity of instruments measuring job satisfaction-a systematic review. *Occupational Medicine*, 53(3), 191–200. <https://doi.org/10.1093/occmed/kqg038>
- Veetkazhi, R., Kamalanabhan, T. J., Malhotra, P., Arora, R., & Mueller, A. (2020). Unethical employee behaviour: A review and typology. *International Journal of Human Resource Management*, 33(10), 1976–2018. <https://doi.org/10.1080/09585192.2020.1810738>
- Wagge, J. R., Brandt, M. J., Lazarevic, L. B., Legate, N., Christopherson, C., Wiggins, B., & Grahe, J. E. (2019). Publishing research with undergraduate students via replication work: The collaborative replications and education project. *Frontiers in Psychology*, 10, 247. <https://doi.org/10.3389/fpsyg.2019.00247>
- Weber, J. (1992). Scenarios in business ethics research: Review, critical assessment, and recommendations. *Business Ethics Quarterly*, 2(2), 137–160. <https://doi.org/10.2307/3857568>
- Weber, J., & Gillespie, J. (1998). Differences in ethical beliefs, intentions, and behaviors: The role of beliefs and intentions in ethics research revisited. *Business & Society*, 37(4), 447–467. <https://doi.org/10.1177/000765039803700406>
- Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, 40(1), 1–29. <https://doi.org/10.18637/jss.v040.i01>
- Wickham, H., François, R., Henry, L., & Müller, K. (2020). *dplyr: A grammar of data Manipulation*. R package version 1.0.2. <https://CRAN.R-project.org/package=dplyr>
- Wiernik, B. M., & Ones, D. S. (2018). Ethical employee behaviors in the consensus taxonomy of counterproductive work behaviors. *International Journal of Selection and Assessment*, 26(1), 36–48. <https://doi.org/10.1111/ijasa.12199>
- Wouters, K., Maesschalck, J., Peeters, C. F., & Roosen, M. (2014). Methodological issues in the design of online surveys for measuring unethical work behavior: Recommendations on the basis of a split-ballot experiment. *Journal of Business Ethics*, 120(2), 275–289. <https://doi.org/10.1007/s10551-013-1659-5>
- Yang, J., Ming, X., Wang, Z., & Adams, S. M. (2017). Are sex effects on ethical decision-making fake or real? A meta-analysis on the contaminating role of social desirability response bias. *Psychological Reports*, 120(1), 25–48. <https://doi.org/10.1177/0033294116682945>

Appendices

Appendix A

Bar plots showing the frequency of responses for each manipulation check

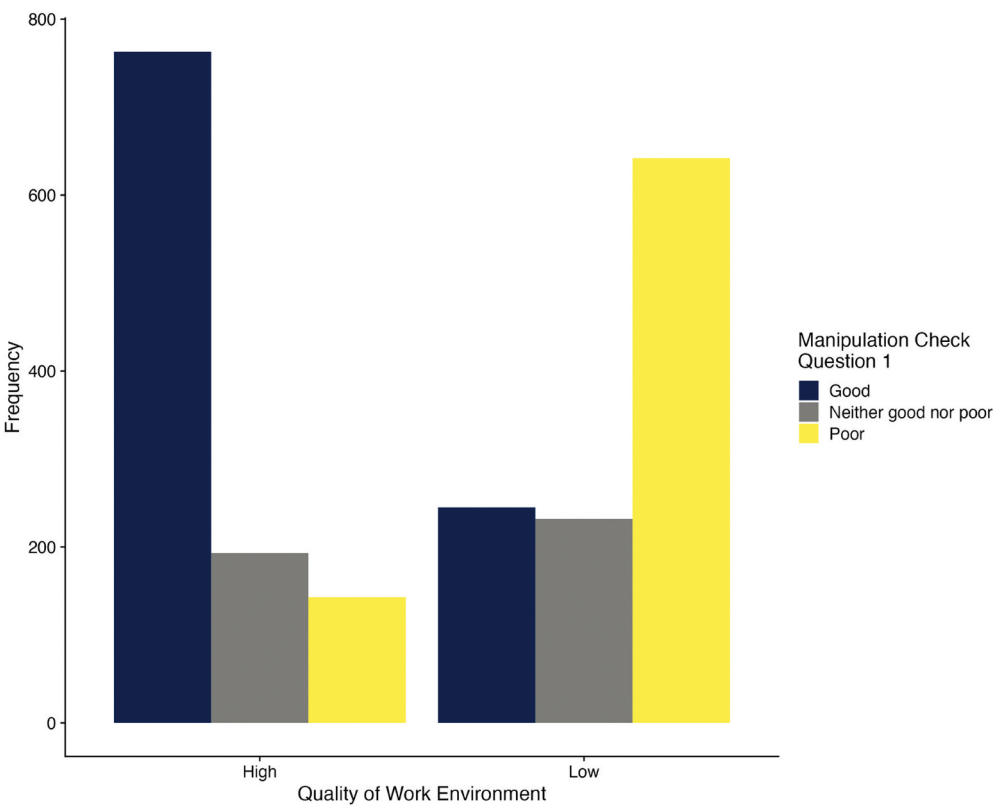


Figure A1. Bar plot for the manipulation check question 1 against the quality of work environment conditions.

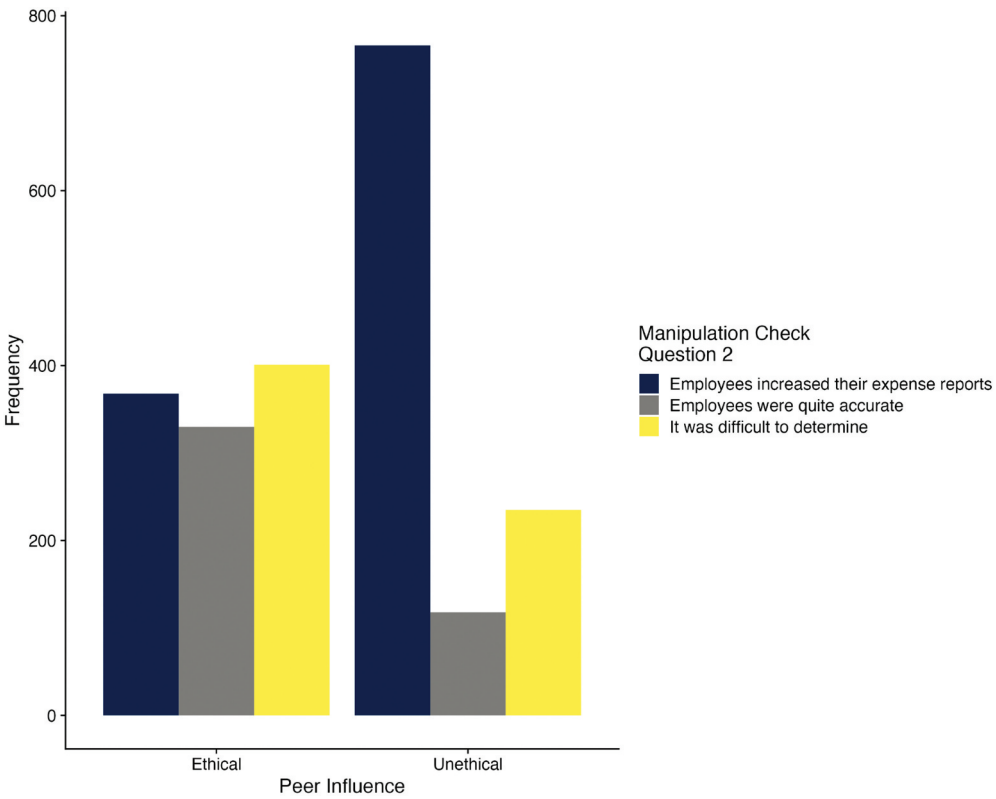


Figure A2. Bar plot for the manipulation check question 2 against the peer influence conditions.

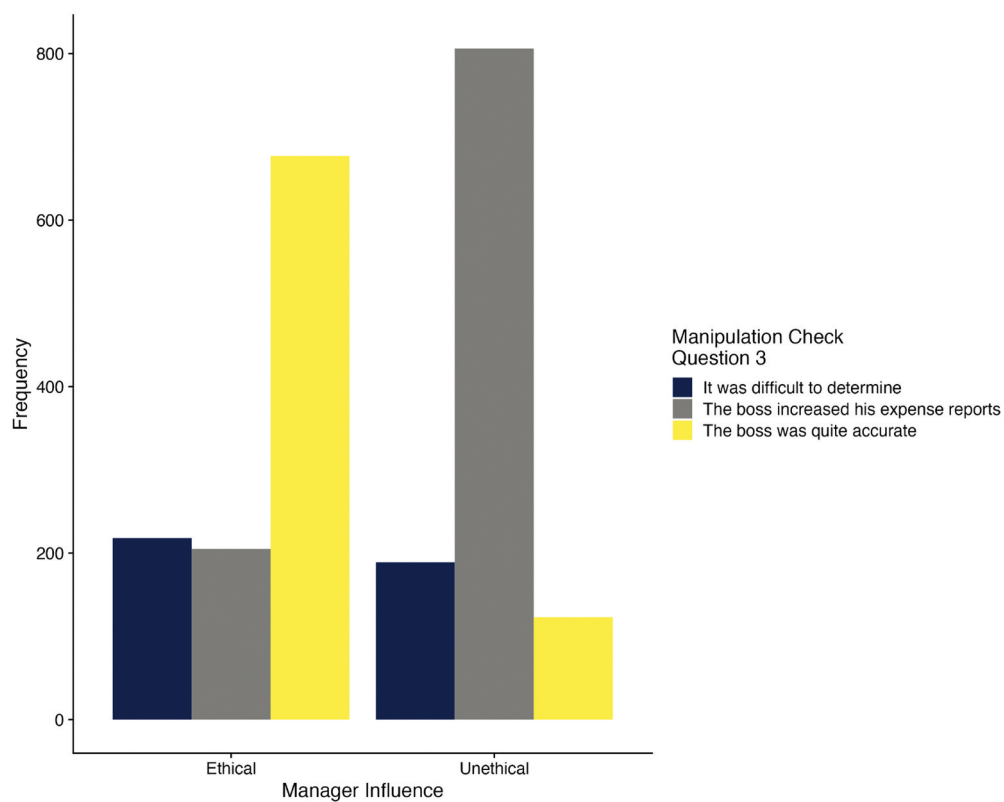


Figure A3. Bar plot for the manipulation check question 3 against the manager influence conditions.

Appendix B

Revised Effect Estimates, including Data Collection Team as Random Effect

Predictor	Direct Replication		Conceptual Replication	
	<i>b</i>	<i>b</i> 95% CI [LL, UL]	<i>b</i>	<i>b</i> 95% CI [LL, UL]
(Intercept)	2.25	[3.07, 5.44]	11.44	[5.60, 17.34]
Quality of Work Experience	−0.41	[−0.70, −0.11]	0.03	[−0.02, 0.08]
Manager influence	0.52	[0.23, 0.82]	−0.04	[−0.09, 0.01]
Peer Influence	0.35	[0.05, 0.64]	0.51	[0.41, 0.61]
Locus of Control	0.04	[0.00, 0.07]	−0.02	[−0.09, 0.04]
Machiavellianism	0.18	[0.15, 0.22]	0.26	[0.20, 0.31]
Gender	−0.61	[−0.95, −0.27]	−3.14	[−4.23, −2.04]