Leadership, creativity, and innovation: A critical review and practical recommendations

David J. Hughes⁎, Allan Leeb, Amy Wei Tianc, Alex Newmanb, Alison Legoodc

⁎ Corresponding author.
E-mail addresses: david.hughes-4@manchester.ac.uk (D.J. Hughes), allan.lee@exeter.ac.uk (A. Lee), Amy.tian@curtin.edu.au (A.W. Tian), a.newman@deakin.edu.au (A. Newman), a.legood2@aston.ac.uk (A. Legood).

Contents lists available at ScienceDirect
The Leadership Quarterly
journal homepage: www.elsevier.com/locate/leaqua

ARTICLE INFO

Keywords:
Leadership
Creativity
Innovation
Measurement
Endogeneity
Mediation

ABSTRACT

Leadership is a key predictor of employee, team, and organizational creativity and innovation. Research in this area holds great promise for the development of intriguing theory and impactful policy implications, but only if empirical studies are conducted rigorously. In the current paper, we report a comprehensive review of a large number of empirical studies (N = 196) exploring leadership and workplace creativity and innovation. Using this article cache, we conducted a number of systematic analyses and built narrative arguments documenting observed trends in five areas. First, we review and offer improved definitions of creativity and innovation. Second, we conduct a systematic review of the main effects of leadership upon creativity and innovation and the variables assumed to moderate these effects. Third, we conduct a systematic review of mediating variables. Fourth, we examine whether the study designs commonly employed are suitable to estimate the causal models central to the field. Fifth, we conduct a critical review of the creativity and innovation measures used, noting that most are sub-optimal. Within these sections, we present a number of taxonomies that organize extant research, highlight understudied areas, and serve as a guide for future variable selection. We conclude by highlighting key suggestions for future research that we hope will reorient the field and improve the rigour of future research such that we can build more reliable and useful theories and policy recommendations.

Introduction

“Creativity, as has been said, consists largely of rearranging what we know in order to find out what we do not know. Hence, to think creatively, we must be able to look afresh at what we normally take for granted.”

George Kneller

Creativity and innovation drive progress and allow organizations to maintain competitive advantage (Anderson, De Dreu, & Nijstad, 2004; Zhou & Shalley, 2003). In recent years, both industry and academia have placed a premium upon creativity and innovation, and research in the field has burgeoned, generating a number of compelling findings (Anderson, Potočnik, & Zhou, 2014). Unfortunately, the research has also been piecemeal in nature. As a result, the leadership, creativity and innovation literature is fragmented and primarily populated by small, ‘exploratory’ studies, which are unrelated to any unifying framework(s). In addition, the rapid growth of research in this field appears to have reduced consideration for a number of fundamental concerns, such as the measurement of key constructs (i.e., creativity and innovation) and the use of study designs that are suitable to address the fascinating research questions posed.

Although leadership has been routinely covered within past reviews of creativity and innovation, it is usually covered briefly, in a descriptive manner, or noted as an area for future research (Anderson et al., 2004, 2014; Rank, Pace, & Free, 2004; Zhou & Shalley, 2003). Previous reviews which have focussed explicitly on leadership and creativity or innovation have typically summarized existing research, provided overviews of dominant theoretical frameworks, identified ‘gaps’ within the literature, and noted practical implications (Klijn & Tomic, 2010; Shalley & Gilson, 2004).

In contrast, our goal is two-fold. First, we aim to summarize the main trends across the myriad of leader variables, mediators and moderators identified within the literature. In doing so, we present a number of taxonomies that synthesize extant research and can guide...
future variable selection, moving studies away from pure exploration toward a more systematic approach. Second, we consider the robustness with which the literature has proceeded so far and draw attention to two major limitations that currently undermine the veracity of the field: measurement and study design. We provide pragmatic guidance so that future research can move beyond these limitations, because left unchecked they stand to limit the scientific and practical merit of research concerning leadership, creativity, and innovation. The nature of our goals in conjunction with the vast array of variables examined in a piecemeal manner and concerns regarding the robustness of many primary studies preclude the use of meta-analytic techniques. Instead, we utilize a combination of systematic and narrative techniques to review the literature. We hope that the recommendations made will help to reorient the field such that future findings will be more robust and generate meaningful policy implications. In essence, we follow the opening quote and hope that by looking afresh at what we normally take for granted, we can help advance research in this vital area.

The remainder of this review is organized as follows. Next, we outline the systematic search strategy that we utilized to identify all papers that had examined leadership and either or both of creativity and innovation. Then we move onto our five substantive review sections. Section 1 revisits a well-trodden path, the conceptualization and definition of creativity and innovation. We aim to make explicit how the two relate and what makes them unique, because, although previous papers have covered this issue, our review suggests that researchers remain unclear. Section 2 provides a systematic review of the leader variables examined and their relationship with creativity and innovation, along with a review and categorization of the proposed moderators of this relationship. Section 3 examines the mediating mechanisms by which leaders are theorized to influence workplace creativity and innovation. Within Section 3, we provide a theoretically-driven taxonomy of these mediating variables, which can be used to guide future research. Section 4 examines the study designs commonly employed, with a particular focus on endogeneity-based concerns. Most often, researchers wish to examine causal process models, whereby leader behavior influences creativity and innovation through some mediating mechanism. Unfortunately, the most frequently employed study designs are not well-suited to assessing such models and making causal inferences. We provide guidance on how researchers can examine such effects in a robust manner. In Section 5, we examine current approaches to measuring creativity and innovation, including an expert review of popular psychometric scales, with a view to establishing what exactly they do and do not measure. Finally, we identify key areas for future research that should produce a more reliable and systematic body of evidence to serve as a platform for theory development and trustworthy policy recommendations.

Search strategy

To review the current empirical literature, we first conducted a comprehensive search for relevant studies. Accordingly, using four databases (Proquest, PsychInfo, EBSCO, and ISI Web of Science) we searched for the keywords “Leadership,” “Leader,” and “Creativity,” “Innovation,” “Creative Behavior,” “Innovative Behavior”. The search included journal articles, dissertations, books chapters, and conference proceedings. We also searched the reference lists from relevant review articles (Anderson et al., 2014; Mainemelis, Kark, & Epitropaki, 2015; Reiter-Palmon & Illies, 2004; Wang, Oh, Courtright, & Colbert, 2011; Zhou & Shalley, 2003).

In total, we identified 185 publications and 195 independent samples (several publications reported multiple samples). Fifty-nine samples were at the team- or organizational-level of analysis, with the remainder being at the individual level. The vast majority of studies used a field sample of employees, and eight studies used a student sample. Throughout this review, we used this article cache to conduct a number of systematic analyses (i.e., documenting all mediators of the leader-creativity/innovation pathway studied) and also as the basis for a number of narrative arguments based on trends evident with these papers. Given the nature of these papers, the majority of our discussion relates to individual employee creativity and innovation, but the overwhelming majority of the points made apply to all levels of analysis.

Section 1: defining creativity and innovation

Creativity and innovation are nuanced concepts that each incorporate a number of distinct but closely related processes that result in distinct but often closely related outcomes (Anderson et al., 2004, 2014). Given the complex and dynamic nature of both creativity and innovation (Mumford & Mcintosh, 2017), it is perhaps unsurprising that they have proven difficult to define and measure (Batey, 2012). Numerous previous reviews have discussed definitional confusion and the limitations it engenders, with most making some recommendations to provide definitional clarity. Perhaps the most notable recent example is Anderson et al.’s (2014, p.1298) review, in which they put forward the following definition of workplace creativity and innovation:

Creativity and innovation at work are the process, outcomes, and products of attempts to develop and introduce new and improved ways of doing things. The creativity stage of this process refers to idea generation, and innovation to the subsequent stage of implementing ideas toward better procedures, practices, or products. Creativity and innovation [...] will invariably result in identifiable benefits.

There is much to admire in the above definition, most notably, it clearly delineates and integrates creativity and innovation. However, it also suffers from a major limitation; it defines creativity and innovation by their outcomes and products. Definitions that draw upon antecedents and outcomes are common in psychological and managerial research, but such definitions are limited for two main reasons (MacKenzie, 2003). First, they do not describe the nature of the phenomenon and thus can lead to misconceptions which, as we discuss later, foster poor measure development (Hughes, 2018; MacKenzie, 2003). Second, they make it difficult (perhaps impossible) to differentiate the phenomenon from its effects: a good joke elicits laughter from an audience, but a joke is still a joke regardless of whether people laugh. The same is true of creativity and innovation, yet the Anderson et al. definition (and many others) states that creativity and innovation are “outcomes and products” that will “invariably [i.e., on every occasion] result in identifiable benefits”. If we follow this logically, an idea cannot be creative until it leads to identifiable benefits to the organization. Even if we leave aside potential concerns regarding the precise meaning of ‘identifiable’, ‘benefits’, and ‘organization’ here, such definitions remain problematic. A creative idea or innovative process cannot exist until after the effects are known – would it really be the case that cars, vaccines, or computers would be considered lacking in creativity if they had not resulted in profitable endeavours? Are we to regard the processes that led to the discovery of DNA as more creative and innovative with each new identifiable benefit we find? Further, such a definition means that creativity and innovation only exist within a particular temporal space. In other words, something can change from being uncreative to creative and back to uncreative again dependent upon market forces; the high-speed aeroplane, Concorde, for example. Clearly, defining creativity and innovation at work by the nature of the effect they have is unhelpful (MacKenzie, 2003).

In a bid to provide unambiguous and succinct definitions that avoid the concerns noted above, yet remain consistent with prior research, we coded every definition provided within our article sample, to identify the core conceptual commonalities while also identifying which are suitable or not as elements of a construct definition (MacKenzie, 2003). An overview of the results of the coding procedure is displayed in Table 1.

In all, 79% of articles provided an explicit definition of either or both creativity and/or innovation. Of those, 47% focused solely on...
Distinguishing between creativity and innovation.

Table 1
Conceptual markers used when defining workplace creativity and innovation.

<table>
<thead>
<tr>
<th>Conceptual properties of creativity definitions (N = 96)</th>
<th>Conceptual properties of innovation definitions (N = 68)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generation of new/novel/original ideas</td>
<td>Problem recognition</td>
</tr>
<tr>
<td>Generation of useful/applicable ideas</td>
<td>Create new ideas/products/processes to be implemented, etc.</td>
</tr>
<tr>
<td></td>
<td>Modify or adopt creative ideas</td>
</tr>
<tr>
<td></td>
<td>Promoting/championing ideas</td>
</tr>
<tr>
<td></td>
<td>Implementation or application</td>
</tr>
<tr>
<td></td>
<td>Organizational benefit</td>
</tr>
</tbody>
</table>

Note: % = the percentage of articles within our cache that defined creativity or innovation with this property.

It is important to note that these definitions conceptualize creativity and innovation independently of any antecedents (e.g., a specific problem) or potential effects (e.g., organizational benefits). In other words, creativity and innovation are not only triggered in certain circumstances and whether generating and implementing ideas leads to improved organizational outcomes is a not a feature of either creativity or innovation, rather it is an outcome. In addition, these integrative definitions, state clearly that creativity and innovation at work are two distinct but closely related concepts, with creativity referring to the generation of novel ideas and innovation referring to (subsequent) efforts to introduce, modify, promote and implement those ideas.

Our definitions are contrary to some, which see innovation as a broad construct that subsumes creativity (e.g., West & Farr, 1990). Although we agree that most innovation starts with a novel idea; arguing that creativity and innovation are synonymous or that creativity can only exist as part of an innovative process is incorrect. Not all creative ideas are taken through the implementation process and not all innovative processes require a creativity (e.g., an organization can innovate by using a non-novel idea taken from elsewhere). In a bid to provide further clarity, we have compiled a comparative table, which provides a succinct and explicit exposition of the differences between creativity and innovation in some key areas (see Table 2).

As is clear from the definitions generated and the characteristics presented in Table 2, creativity (idea generation) and innovation (implementation) are different constructs that arise as the result of distinct processes and lead to different outcomes. Thus, we can distinguish between the two conceptually, and so, we should be able to distinguish between them empirically. However, our review suggests that, too often, researchers treat the two as synonyms with authors citing creativity research to build hypotheses related to innovation and vice versa (e.g., Kao, Pai, Lin, & Zhong, 2015; Zhu, Wang, Zheng, Liu, & Miao, 2013). Further, a number of papers, even when published in top-tier journals, discussed creativity but used scales that purport to assess innovation (e.g., Neubert, Kacmar, Carlson, Chonko, & Roberts, 2008; Zhang, Tsui, & Wang, 2011) and vice versa (e.g., Zhu et al., 2013). Despite numerous warnings (e.g., Rank et al., 2004), researchers have failed to heed the nuance here: yes, creativity and innovation are related, but “they are by no means identical” (Anderson et al., 2014, p. 1299).

Indeed, where efforts have been made to provide nuanced measurement, evidence suggests that creativity and innovation have different antecedents. For example, individual-level variables such as self-efficacy are predictors of idea generation, whereas managerial support is a predictor of innovative endeavours (e.g., Axtell et al., 2000; Magadley & Birdi, 2012). Further, it is not inconceivable that the different elements of creativity and innovation have different antecedents; it would not be surprising if idea generation was most closely associated to the personality trait of openness, that idea promotion was most closely associated to extraversion, and that idea implementation was most closely associated to conscientiousness. Clearly, the parsing of the
Table 3
Key leadership variables studied, along with definitions and main study characteristics including the study design and the source of the creativity/innovation rating.

<table>
<thead>
<tr>
<th>Leadership variable</th>
<th>Definition</th>
<th>Study characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Study characteristics</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Creativity (self-rated</td>
<td>other-rated)</td>
</tr>
<tr>
<td></td>
<td>XS</td>
<td>TS</td>
</tr>
<tr>
<td>Transformational leadership (Bass, 1985)</td>
<td>Theorized to consist of four dimensions. First, idealized influence reflects the degree to which the leader behaves admirably and causes followers to identify with the leader. Second, inspirational motivation reflects the degree to which the leader articulates an appealing and inspiring vision. Third, intellectual stimulation reflects the degree to which the leader challenges assumptions, takes risks, and solicits followers' ideas. Fourth, individualized consideration reflects the degree to which the leader listens and attends to each follower's needs, and acts as a mentor or coach.</td>
<td>7</td>
</tr>
<tr>
<td>Leader-member exchange (LMX; Dansereau, Graen, &amp; Haga, 1975)</td>
<td>LMX theory posits that, through various exchanges, leaders differentiate in the way they treat followers, leading to different quality relationships. Research shows that high LMX quality is associated with a range of positive follower outcomes</td>
<td>3</td>
</tr>
<tr>
<td>Transactional leadership (Bass, 1985)</td>
<td>Transactional leadership has three dimensions: contingent reward, management by exception—active, and management by exception—passive. Contingent reward reflects the degree to which the leader enacts constructive transactions or exchanges with followers. Management by exception is the degree to which the leader takes corrective action on the basis of results of leader-follower transactions.</td>
<td>2</td>
</tr>
<tr>
<td>Empowering leadership (Kirkman &amp; Rosen, 1999)</td>
<td>Empowering leadership entails delegation of authority to employees, promotion of self-directed and autonomous decision-making, coaching, sharing information, and asking for input.</td>
<td>4</td>
</tr>
<tr>
<td>Authentic leadership (Walumbwa, Avolio, Gardner, Wemming, &amp; Peterson, 2008)</td>
<td>Authentic leadership builds upon and promotes positive psychological capacities and a constructive ethical climate, to foster greater self-awareness, an internalized moral perspective, balanced processing of information, and relational transparency between leaders and followers.</td>
<td>2</td>
</tr>
<tr>
<td>Servant Leadership (Hale &amp; Fields, 2007)</td>
<td>Servant leadership places the interests of followers over the self-interest of the leader, emphasizing leader behaviors that focus on follower development, and de-emphasizing elevation of the leader.</td>
<td>3</td>
</tr>
<tr>
<td>Totals</td>
<td>Aggregate number of papers in each study design category summed using all papers using any leadership style</td>
<td>21</td>
</tr>
</tbody>
</table>

Note: XS = cross-sectional; TS = time separated; L = longitudinal; EX = experimental; numbers to the left of the | represent studies with self-ratings of creativity/innovation and numbers to the right of the | represent studies with other-ratings of creativity/innovation.
two constructs holds potential for the development of more nuanced and useful theories, empirical estimates, and practical implications.

We make one simple recommendation at this point: researchers must be precise in their use of terminology and existing literature. Leadership, creativity, and innovation research does not advance when we conflate and confuse constructs. For now, we leave the discussion of conceptualizing creativity and innovation, but we return to it in the measurement section of our review. As discussed there, the lack of care regarding the conceptualization of the two constructs has had a negative effect on the quality of measurement tools available to researchers, which is one of the biggest factors preventing the field from fulfilling its potential. Thus, we need to develop new tools that provide accurate and appropriate measurement (Hughes, 2018) of these two constructs.

Section 2: leadership

Many leadership variables have been examined as predictors of workplace creativity and innovation. We have compiled two tables to provide a broad, descriptive summary of this literature. Table 3 contains descriptions and definitions of the most commonly studied leadership variables, a breakdown of the number of studies investigating them, and we note the major study design employed (i.e., cross-sectional versus experimental) which we will discuss later. Table 3 reveals that the most studied leadership approaches are the well-established, transformational leadership (N = 81) and leader-member-exchange (LMX, N = 48). In contrast, newer approaches, such as empowering, servant, and authentic leadership, have received less attention. An interesting observation is that most assessments of leaders have focused on ‘leader styles’, with leader traits, such as personality and IQ, receiving very little attention. The omission of well-established individual differences is surprising, because studies that have measured such leader characteristics have found significant associations with follower creativity (e.g., Huang, Krasikova, & Liu, 2016).

Table 4 summarizes the range and average strength of associations between frequently studied leadership approaches and both creativity and innovation. There are two broad trends evident within Table 4. First, and perhaps unsurprisingly, leadership styles typically considered ‘constructive’ or ‘positive’ (e.g., transformational, empowering) are positively associated with both creativity and innovation. Second, within each leadership style, there is a large degree of variability in observed associations. We now explore these points in relation to specific leadership styles.

Transformational and transactional leadership are perhaps the best known within and outside the leadership field, and the two are often pitted as opposite or competing approaches. However, our review of research using these variables has provided some interesting and somewhat analogous findings, which provide some general points for the literature en masse. Both have small, average positive correlations with creativity and innovation and also demonstrate the largest range of observed correlations. We believe a number of factors have produced this pattern of findings.

First, transformational (N = 75) and transactional (N = 16) have been included within a large number of studies and so the variation might represent differences in samples, contexts, measurement tools, and rating-sources. Future meta-analytic studies focussed on uncovering the extent to which the context (e.g., industry, role) and study-design have served to moderate the effects observed would help determine whether this is the case. It is also possible that the instability or low reliability of the findings is a product of endogeneity biases that result from sub-optimal study design, which we discuss in more detail in Section 4: study design.

Second, both transactional and transformational leadership consist of several lower-order factors or components (see Table 3). However, studies have tended to operationalize these leadership styles through a single scale score, thus ignoring and masking any sub-factor-level relationships (e.g., Miao, Newman, & Lamb, 2012), which is a pertinent limitation.

With regard to transformational leadership, theory suggests that some sub-factors might be more relevant than others. For example, intellectual stimulation and inspirational motivation have been specifically highlighted as critical for innovation (Elkins & Keller, 2003). For instance, by providing intellectual stimulation (Bass & Avolio, 1997), leaders can encourage followers to adopt generative and exploratory thinking processes (Sosik, Avolio, & Kahai, 1997) that are likely to support and stimulate employees to contend with unusual challenges and problems (Srivastava, Bartol, & Locke, 2006). Although we see great value in nuanced sub-factor examinations, we must also note that recent research has provided some compelling critiques regarding the multi-dimensional definition of transformational leadership, focusing on theoretical ambiguities, the insufficient specification of causal processes (Van Knippenberg & Sitkin, 2013), and problems with popular measures (i.e., Multifactor Leadership Questionnaire: Antonakis, Bastardoz, Jacquart, & Shamir, 2016; Van Knippenberg & Sitkin, 2013).

In contrast to encouraging intellectual stimulation and providing inspirational motivation, transactional leadership describes leader behaviors that utilize extrinsic motivation in something of a ‘quid pro quo’ style: if followers perform well, they are rewarded (contingent reward), if not, they are reprimanded (management by exception). In general, because transactional leadership does not engender intrinsic motivation, it is considered to stifle creativity and innovation (e.g., Amabile,

<table>
<thead>
<tr>
<th>Leadership approach</th>
<th>Creativity</th>
<th>Innovation</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Range</td>
<td>Average</td>
</tr>
<tr>
<td>Transformational (overall)</td>
<td>36</td>
<td>−0.13–0.68</td>
</tr>
<tr>
<td>Idealized Influence</td>
<td>4</td>
<td>0.08–0.18</td>
</tr>
<tr>
<td>Inspirational motivation</td>
<td>4</td>
<td>0.11–0.21</td>
</tr>
<tr>
<td>Intellectual stimulation</td>
<td>3</td>
<td>0.05–0.19</td>
</tr>
<tr>
<td>Individualized consideration</td>
<td>5</td>
<td>0.08–0.27</td>
</tr>
<tr>
<td>LMX</td>
<td>22</td>
<td>0.04–0.65</td>
</tr>
<tr>
<td>Transactional (overall)</td>
<td>6</td>
<td>−0.29–0.46</td>
</tr>
<tr>
<td>Contingent reward</td>
<td>2</td>
<td>0.12–0.14</td>
</tr>
<tr>
<td>Active management by exception</td>
<td>1</td>
<td>−0.03</td>
</tr>
<tr>
<td>Passive management by exception</td>
<td>2</td>
<td>−0.06–0.02</td>
</tr>
<tr>
<td>Empowering leadership</td>
<td>15</td>
<td>0.20–0.66</td>
</tr>
<tr>
<td>Authentic leadership</td>
<td>7</td>
<td>0.01–0.75</td>
</tr>
<tr>
<td>Servant leadership</td>
<td>8</td>
<td>−0.04–0.59</td>
</tr>
</tbody>
</table>

Note: The column ‘Average’ indicates the magnitude of the average correlation based on Cohen’s (1992) rule of thumb; − = average correlation is ≤ 0.10; + = (small) average r is between 0.10 and 0.30; ++ = (medium) average r is between 0.30 and 0.50. The studies used to calculate the range and average effect sizes are marked with an * in the reference list.
1996). However, the sub-factor of contingent reward sometimes correlates positively with both creative (e.g., $r = 0.46$; Rickards, Chen, & Moger, 2001) and innovative behavior (e.g., $r = 0.58$, Chang, Bai, & Li, 2015). Indeed, a recent meta-analysis suggested rewards contingent upon employee creativity rather than performance or task completion are particularly effective (Byron & Khazanchi, 2012). In contrast, the management by exception sub-factor consistently correlates negatively with creativity and innovation (Moss & Ritossa, 2007; Rank, Nelson, Allen, & Xu, 2009).

Thus, by failing to explore the sub-factors of transformational and transactional leadership, it is possible, perhaps probable, that our current estimates of the effects of these leadership styles are sub-optimal. Not only is it likely that certain sub-factors may be more predictive, but, as alluded to above and discussed in greater detail below, they might also speak to different mediators (e.g., Van Knippenberg & Sitkin, 2013). To build more comprehensive models of the leader-creativity/innovation relationship, we urge future research to examine these sub-factors and perhaps consider exploring the full-range leadership model (Bass & Avolio, 1995) or better still the “fuller full-range” leadership model (Antonakis & House, 2014).

As noted above, LMX, a relational approach to leadership (see Table 3), is the second most frequently studied leadership variable. In leader-follower relationships characterized by high levels of LMX quality, leaders may stimulate creative and innovative performance by providing followers with high levels of autonomy and discretion (e.g., Pan, Sun, & Chow, 2012), allocating needed resources (e.g., Gu, Tang, & Jiang, 2015), and building followers’ confidence (Liao, Liu, & Loi, 2010). Most often, LMX was used as leadership-based predictor modelled as having either a direct (e.g., Lee, 2008) or indirect effect (e.g., Liao et al., 2010) on creativity or innovation, but it was also used as a mediator (Gu et al., 2015) and moderator (Van Dyne, Jhehn, & Cummings, 2002). Typically, results in these studies support the hypothesized effect, whether a main effect, mediation or moderation. As such, our review reflects an interesting plurality that exists within the wider leadership literature regarding the theoretical status of LMX (e.g., Lee, Willis, & Tian, 2017). Clearly, the lack of conceptual clarity regarding LMX is an issue that needs to be addressed. In addition to the theoretical dilemmas, there are also potentially statistical concerns associated with employing LMX as a predictor of creativity or innovation (Fairhurst & Antonakis, 2012; House & Aditya, 1997). Briefly, because LMX refers to a rating of relationship quality between leader and follower, it is technically speaking the outcome of a leader behavior-follower reaction process. Thus, LMX is an outcome variable in and of itself and, when employed as a predictor, we are essentially relating one outcome to another. Statistically speaking, this means that LMX is an endogenous variable, and endogenous predictors are associated with several biases that can influence the reliability and veracity of parameter estimates. We discuss endogenous variables in more detail in Section 4: study design.

Given the criticisms and conceptual issues associated with some of the more well-established leadership theories (e.g., Antonakis et al., 2016; Van Knippenberg & Sitkin, 2013), it is not surprising that our review revealed that examinations of contemporary leadership styles, such as empowering, servant and authentic leadership, have recently increased. Research examining these contemporary styles suggests that they have a relatively strong association with both outcomes (Table 4). In particular, empowering and authentic leadership have moderate average associations and smaller ranges than the others. In addition, and not covered in Table 4, are two recent studies that suggest that the negative association between aversive leadership and creativity and innovation are stronger than the positive associations of positive leadership. Specifically, despotic (Naseer, Raja, Syed, Donia, & Darr, 2016) and authoritarian leadership styles (Wang, Chiang, Tsai, Lin, & Cheng, 2013) showed stronger associations with creativity than LMX or benevolent leadership, respectively.

Based on our review, it seems that more contemporary and narrowly specified leader variables tend to have larger effects than do broad measures of well-established leader variables. However, currently the evidence needed to make definitive conclusions is not available. Specifically, few studies have used appropriate designs to examine multiple leader variables concurrently. Thus, we have little direct evidence regarding the relative or incremental predictive effects of these many different leader variables. The few studies that have examined relative or incremental effects suggest that different leader variables do not contribute equally. For instance, whereas some studies showed that transactional leadership had a significant negative association with innovation when examined alongside transformational leadership (Lee, 2008; Pieterse, van Knippenberg, Schippers, & Stam, 2010), McMurray, Islam, Sarros, and Pirola-Merlo (2013) found that the use of contingent punishments had stronger positive association with innovation than did transformational leadership. In addition, a number of studies suggested that LMX shared stronger associations with creativity and innovation than transformational leadership (e.g., Pundt, 2015; Turunc, Celik, Tabak, & Kabak, 2010), contingent rewards (Turunc et al., 2010), and humorous leadership (Pundt, 2015).

As is clear from this review, many leader variables share roughly equivalent associations with follower creativity and innovation and, as discussed shortly, are also theorized to influence creativity and innovation through the same mediating mechanisms. Although interesting, the observed homogeneity is likely a reflection of construct proliferation and construct redundancy within leadership research (Shaffer, DeGeest, & Li, 2016), which has produced an overly complex literature that hinders understanding, theory building, and the development of practical recommendations (Derue, Nahrgang, Wellman, & Humphrey, 2011; Shaffer et al., 2016). Thus, it is important that future studies make a concerted effort, using appropriate study designs (i.e., experimental or instrumental variable, see Section 4: study design), to address the relative and incremental effects of different leader variables and identify which parsimonious combination of leadership variables best fosters creativity and innovation. The few studies which have examined multiple leader variables suggest that such endeavours are likely to be fruitful.

**Moderating variables**

The magnitude of the relationship between leadership and creativity and innovation is hugely variable (see Table 4). In some cases, ranging from near-zero to large, and in others, ranging from moderately negative to moderately positive. There are three likely contributing factors to this variation. First, the variation could be a methodological artefact resulting from differences in study design. For example, some studies are experimental in nature (e.g., Boies, Fiset, & Gill, 2015) but many more are survey-based field studies (e.g., Zhang & Bartol, 2010). As discussed later, the latter are particularly susceptible to a range of measurement-based limitations. Second, the variation might represent the fact that the very nature of creativity and innovation differs across organizational sectors and roles. For example, it is possible that the leadership needed to support innovation in sales industries differs from manufacturing industries. Currently, no papers have empirically examined cross-industry effects, thus, direct comparisons across industry boundaries would be an interesting avenue for future research. Third, the variation might reflect the presence of moderating, within-context variables that influence the nature of the relationships.

In recent years, studies ($N = 36$) have investigated a wide-range of moderating variables that exacerbate and attenuate the positive effects of “positive leadership” and the negative effects of “negative leadership.” The moderators can be categorized as attributes of the follower (e.g., personality, motivation), the leader (e.g., gender, encouragement of creativity), the leader-follower relationship (e.g., LMX, identification with the leader), or aspects of the team or organizational context (e.g., organizational structure, team relational conflict). See Fig. 1 for a summary.
The broad range of moderators studied makes a succinct summary difficult, especially given the approach to exploring moderation tends to utilize idiosyncratic, micro-theoretical, study-specific reasoning for hypothesis development. In other words, the array of moderators investigated lacks any coherent theoretical narrative and most have been examined just once. In future, research would benefit greatly from a unifying theoretical framework or even taxonomic classification. In Fig. 1, we have provided a data-driven taxonomy, which, in the absence of a theoretical framework, can be useful. We would urge researchers to: (i) justify the category (e.g., leader attributes or contextual attributes) they have chosen to explore, and (ii) justify why their chosen moderator is more appropriate than other moderators within their chosen category. In addition to providing a more thorough rationale, future studies must use study designs that are robust to endogeneity bias to accurately estimate moderation effects (see Section 4: study design for further discussion). Overall, it is positive that researchers are beginning to examine the conditions that render various leadership approaches more or less effective, but future research would benefit greatly from a clear theoretical framework and more rigorous study designs.

Section 3: mediating mechanisms

Leadership is a process whereby leader variables affect distal outcomes (i.e., creativity and innovation) through more proximate mediating variables (e.g., follower motivation: Fischer, Dietz, & Antonakis, 2017). Accordingly, many studies ($N = 64$) within our review examined mediating mechanisms. Examining meditational processes is
integral to the development of theory and practical recommendations. However, our review revealed two notable limitations. First, the study designs commonly used are sub-optimal, due to their susceptibility to endogeneity biases, which we discuss further in the next section. Second and our major focus here is the unsystematic approach taken to the selection of mediators. Different leadership approaches should proffer distinct theoretical explanations of the mechanisms through which they influence followers’ creative or innovative behavior. However, there is a great deal of overlap across many approaches. Thus, within this review we aim to provide a comprehensive but parsimonious taxonomy of the mediators commonly explored to help guide future research. Two previous reviews of workplace creativity and innovation have highlighted motivational, cognitive and affective mechanisms which can mediate the effects of leadership on creativity and innovation (Shin, 2015; Zhou & Shalley, 2011). In addition, our systematic review identified two additional mediating mechanisms: identification-based and relational-based. The five classes of mediators, with exhaustive lists of specific variables examined, are depicted in Fig. 2.

We hope that the taxonomy not only describes the extant literature but also guides and refines future work. For example, during study design, researchers should first identify which broad mechanism they suspect is most relevant for their purpose (e.g., motivational or cognitive). Then they should review each variable within their chosen class of mediators and select the most appropriate variable(s). We would also suggest that researchers choose an additional mediator from a different category so that they can provide a more compelling test of the utility and uniqueness of their preferred mediator(s). In addition, we hope that by identifying theoretically distinct classes of mediators we have provided a framework for researchers to identify the mechanisms within and across classes is most closely associated with each leader variable and with different domains of creativity and innovation. In essence, we urge researchers to use this theoretically derived taxonomy to guide variable selection, conduct more rigorous empirical tests, and refine the literature, so that we can move toward more parsimonious, powerful and useful models of leadership, creativity and innovation. Below, we discuss why each of these mediational mechanisms should relate to workplace creativity and innovation, and note areas for future research.

**Motivational mechanisms**

Motivational mechanisms have received the most attention at individual, team and organizational levels. Both Amabile’s (1996) influential componential theory of creativity and Scott and Bruce’s (1994) seminal innovation paper place intrinsic motivation as a key driver of workplace creativity and innovation. Intrinsic motivation results from individuals’ interest and involvement in, satisfaction with, or positive challenge associated with task engagement (e.g., Deci & Ryan, 2000). Intrinsic motivation is particularly important for creativity and innovation, because acts of the two often fall outside of normal work tasks and require employees to challenge accepted practices. Thus, in addition to possessing the relevant skills and knowledge, employees need to be intrinsically motivated to engage in and persist with the task (Amabile, 1996).

Given the centrality of intrinsic motivation to theories of creativity and innovation, it is unsurprising that variables such as intrinsic motivation, psychological empowerment and creative self-efficacy are...
frequently examined. However, many of these mediators are conceptually and empirically overlapping and thus, although this area of research looks well-developed, it is somewhat narrow. For instance, the notion of extrinsic motivation is entirely absent, probably because self-determination theory suggests that the use of rewards to enhance extrinsic motivation tends to reduce intrinsic motivation and self-determination (e.g., Eisenberger & Aselage, 2009). However, as noted previously, with respect to the use of contingent rewards, the evidence is mixed (Byron & Khazanchi, 2012). Within our sample, we found studies showing both positive (e.g., Chang et al., 2015) and negative (e.g., Lee, 2008) associations between use of contingent rewards and creativity and innovation, with a recent meta-analysis demonstrating that creativity-contingent rewards enhanced creative performance, but that performance- or completion-contingent rewards did not (Byron & Khazanchi, 2012). The effect of creativity-contingent rewards was further enhanced when coupled with positive and specific feedback, and when employees were provided choice regarding the nature of the reward (Byron & Khazanchi, 2012). These findings suggest that when used appropriately rewards, and thus extrinsic motives, are likely to promote workplace creativity and innovation. In addition, Henker, Sonnentag, and Unger (2015) found that employees’ promotion focus mediated the effect of transformational leadership on creativity. Promotion focus is one of the two regulatory foci defined in the Regulatory Focus Theory (Higgins, 1997) and is especially relevant for creativity because it is related to eagerness and risk-taking (e.g., Kark & Van Dijk, 2007), which can be drivers of creative exploration and innovative implementation (Henker et al., 2015).

Cognitive mechanisms

Creative performance requires that employees exhibit relevant cognitive skills and engage in extensive and effortful cognitive processes (Amabile, 1996; Shin, 2015). Research on cognitive mechanisms posits that observed differences in creativity and innovation result from differences in individuals’ use of certain cognitive processes, the capacity of memory systems, and the flexibility of stored cognitive structures (e.g., Ward, Smith, & Finke, 1999). Researchers have started to investigate the role that leaders can play in influencing followers’ cognitive processes. For example, by providing access to diverse information, inspiring team members to share knowledge and ideas, or creating an environment conducive to engagement in creative processes (Reiter-Palmon & Illies, 2004; Shin, 2015). In particular, creative process engagement and support for innovation have been studied most frequently.

Affective mechanisms

Positive affect has long been established as an antecedent of creativity, through laboratory experiments, field studies, and diary studies (e.g., Amabile, Barsade, Mueller, & Staw, 2005). However, only a small number of studies, each supportive, have examined positive affect as a mediator of positive leadership styles (e.g., authentic leadership: Rego, Sousa, Marques, & e Cunha, 2014). The limited exploration of affective mediators is a pertinent limitation because numerous theoretical arguments suggest that both positive and negative affect are likely to influence creativity and persistence (e.g., George & Zhou, 2002). Even ambivalent emotions have been argued to foster creativity because the unusual experience associated with emotional ambivalence signals to individuals they are in an unusual environment, which encourages them to draw upon their creative thinking ability (Fong, 2006). It is also interesting to note that although there is a good degree of theorizing regarding affect and creativity, there is little regarding innovation. The few studies that have explored the relationship between affect and innovation show promising links (e.g., Zhou, Ma, Cheng, & Xia, 2014), and so this looks like a fruitful avenue for future research.

Identification mechanisms

Identification-based mediators represent an alternate form of motivational mechanism (e.g., Tierney, 2015) that draw on self-concept theory (Shamir, House, & Arthur, 1993), role identity theories (Burke & Tully, 1977), and the relational identification concept (Cooper & Thatcher, 2010). For example, research suggests that priming subordinates’ identification with their leaders is crucial for leaders to influence their followers’ beliefs and behaviors (Kark, Shamir, & Chen, 2003), though empirical studies examining creativity and innovation have reported inconsistent results. For instance, identification with leader was found to mediate the effects of transformational leadership (Qu, Janssen, & Shi, 2015) and moral leadership (Gu et al., 2015) on employee creativity and innovation, but it did not mediate the effect of transformational leadership on innovation (Miao et al., 2012). In a rare team level study, team identification with leader mediated the effect of servant leadership on team creativity (Yoshida, Sendjaya, Hirst, & Cooper, 2014). Another more frequently studied identification mechanism is employees’ creative role identity (Farmer, Tierney, & King-McIntyre, 2003). For example, two studies tested and supported the mediating effect of creative role identity between transformational leadership and creativity (Wang, Tsai, & Tsai, 2014; Wang & Zhu, 2011). Overall, identification-based mediators have received little attention, but what research there is suggests that these mechanisms are of value in understanding how leadership influences employee creativity and innovation.

Social relational mechanisms

Social-relational mechanisms are built upon the foundation of social-exchange theory (Blau, 1964). As shown in Fig. 2, this class includes variables such as trust-in-the-leader, LMX (i.e., the quality of the leader-follower relationship), and felt obligation (i.e., whether followers perceive a need to reciprocate favorable leader treatment). According to social-exchange theory, positive exchanges between leaders and followers might lead to creativity and innovation, because followers seek to repay favorable leader treatment by engaging in in-role and extra-role performance (e.g., Martin, Thomas, Guillaume, Lee, & Epitropaki, 2016).

Trust-in-the-leader has been found to play a key role in the development and deepening of leader-follower social exchanges because it encourages obligation and reduces uncertainty around reciprocation (Konovsky & Pugh, 1994). In addition, trust is a crucial facilitator of creativity and innovation because of their inherently unpredictable and risky nature (i.e., suggesting novel ideas often faces resistance and intended benefits are often far from guaranteed). Higher levels of trust lessen the perceived risk and create a psychologically safe environment which facilitates employees’ willingness to engage in creative and innovative actions (Zhang & Zhou, 2014).

However, claims that “trust in the leader is the major lynchpin” in leadership-creativity/innovation relationships (Ng & Feldman, 2015, p. 949), are currently inconsistently supported by empirical evidence. For example, Jaiswal and Dhar (2015) found that trust in the leader mediated the association between servant leadership and employee creativity, but Jo, Lee, Lee, and Hahn (2015) did not find support for trust as a mediator between leader consideration, initiating structure, and employee creativity. Similarly, some studies found that LMX mediated the association of moral leadership (Gu et al., 2015) and transformational leadership (Lee, 2008) with employee creativity and innovation. However, others reported that LMX did not mediate the associations of transformational leadership (Turunc et al., 2010) and transactional leadership with innovation (Lee, 2008; Turunc et al., 2010). Further research is needed to understand the extent to which social relational mechanisms explain the effects of leadership on creativity and innovation. It is also interesting to note the potential theoretical tension between LMX and trust in the leader. Some scholars have positioned trust as a defining feature of LMX (Liden & Graen, 1980) and others (Dirks & Ferrin, 2002) have argued (based on meta-analytic evidence) that the two are related but distinct. Thus it is not clear whether they offer distinct or largely overlapping mediating
mechanisms, with recent meta-analytic findings suggesting that trust in the leader mediated the association between both transformational and empowering leadership on creativity, whereas LMX did not (Lee et al., 2017). Regardless of the exact relationship between the two, the theoretical explanations for the relevance of both hinge on social exchange. Future research should aim to add clarity to the literature by continuing to examine this issue.

Overview of mediational processes

The discussion above highlights that a great deal of research effort has gone into trying to elucidate the underlying mechanisms that explain how various leaders influence creativity and innovation. However, several pertinent limitations are evident. Typically, mediation studies assess a single leader variable and a single mediator. Given the conceptual and empirical overlap between leader variables (as discussed above) and many of the mediators examined (e.g., trust and psychological safety, creative self-efficacy and creative identity: Van Knippenberg & Sitkin, 2013), it is likely that the literature suffers from construct proliferation and redundancy (Shafer et al., 2016). Not only that, but the single leader variable, single mediator designs make it impossible to assess which mediators are most important for creativity and innovation, and which leader variables are most important for each mediator. So, for example, psychological empowerment has been found to mediate the effects of transformational leadership (e.g., Afar, Badir, & Bin Saeed, 2014), empowering leadership (Chen, Sharma, Edinger, Shapiro, & Farh, 2011), servant leadership (e.g., Krog & Govender, 2015a, 2015b), and LMX (Pan et al., 2012). In other words, psychological empowerment mediates the effects of almost all leader variables, which is problematic, because theories or bodies of evidence that include everything explain nothing. By examining multiple leader variables and mediators concurrently, we can rule out some of these effects and build a more parsimonious and useful picture of what is going on.

Similarly, few studies have even examined conceptually dissimilar mediators concurrently (i.e., those from different mechanism categories) and thus we cannot say, for example, whether motivational mechanisms or cognitive mechanisms are more powerful, and whether their effects are additive or not. Future research should begin to address which leadership styles or even dimensions of leadership styles fit best with which mechanisms and, subsequently, which mechanisms are more or less important (Shin, 2015). As we noted at the outset of this section, we hope that the five-category taxonomy will aid in these endeavours by providing a broad framework that can be used to refine study designs. Specifically, researchers can easily identify mediating variables that operate through the same or different broad mechanisms. Thus, this taxonomy describes the extant literature and can be used to guide and refine future research.

As discussed above, mediators of the link between leadership and creativity/innovation can be organized into five broad categories and we believe that our categories can be used in conjunction with those recently identified by Fischer et al. (2017). Fischer and colleagues suggested that mediators within leadership process models can be organized into two distinct types: those that develop or leverage resources. In other words, leaders can leverage (or utilize or mobilize) existing resources, such as employee motivation, or they can develop employees through resource-enlarging activities, such as individual- and/or team-level mentoring and coaching. What is evident from our review, summarized in Fig. 2, is that all the mediators previously examined focus on leveraging existing resources. By ignoring the developmental processes, research has created an imbalance in our current understanding. As Amabile (1996) highlights in her componential theory, for individuals to exhibit high levels of creativity, three components must be present: individuals should possess (a) domain-relevant knowledge and skills, (b) creativity-relevant skills and strategies, and (c) they need to be motivated to work on the task. The first two components of this model focus explicitly on the skills and abilities that are a prerequisite for creativity. However, research has ignored the ways in which leaders can enhance such skills, focussing predominantly on the third component (i.e., extracting maximum enthusiasm and motivation from employees). A clear aim for future research is to address this gap. For example, studies might investigate how leaders can develop skills and knowledge through developmental feedback and knowledge-sharing strategies, which should allow them to acquire and use creativity-relevant skills, strategies and knowledge (Zhou, 2003).

Another key limitation is that many of the mediators identified in Fig. 2 represent psychological states, all of which are assessed through questionnaire ratings, meaning that there is a strong possibility that the different measurements share a strong evaluative component (Van Knippenberg & Sitkin, 2013). Thus, if mediators are studied conjointly (as suggested above), they might not emerge as empirically distinct constructs due to common method bias rather than empirical redundancy. This brings us to perhaps the biggest limiting factor in leadership-creativity/innovation research: study design. We have briefly noted study design issues throughout our review, and in the next section, we deal with them explicitly. It is vital that researchers wish to test causal process models, such as those discussed here, they employ appropriate study designs (Antonakis, Bendahan, Jacquart, & Lalive, 2016; Fischer et al., 2017).

Section 4: study design

As outlined above, the underlying assumption guiding leadership-creativity/innovation research is that leaders can, either directly or indirectly, influence (or statistically speaking, cause) increases or decreases in the frequency and quality of the creativity and innovation displayed by their subordinates. As is evident from Table 3, the typical study uses a cross-sectional design (i.e., whereby all the study variables were measured at the same time) and assesses creativity and innovation through self- (N = 58) or other-ratings (usually a manager; N = 73). In total, 80% of studies we identified utilized such a design and many examined causal process models along the lines of leadership → mediator → creativity/innovation. Unfortunately, these designs, without the use of an instrumental variable procedure, which we discuss shortly, are not capable of providing robust estimates of causal effects due to endogeneity biases (Antonakis et al., 2010; Antonakis, Bendahan, Jacquart, & Lalive, 2014; Fischer et al., 2017; Hamilton & Nickerson, 2003).

Briefly, endogeneity refers to an instance when a predictor variable (whether classed as predictor, mediator, or moderator) is correlated with the error term of the outcome variable (see Antonakis et al., 2010, 2014 for details). In other words, an endogenous predictor is related to the measured outcome variable in two or more ways, usually in the way theorized (e.g., as a meaningful cause), but also in some unanticipated way(s) (e.g., common method bias, reciprocal effects, relationship with a common cause).

Consider a typical cross-sectional, dyadic study in which employees rate their perception of their leader’s authenticity (predictor) their own levels of motivation (mediator), and the leader rates the employees’ creativity (outcome). This study is likely afflicted by endogeneity biases in three domains that affect both leader authenticity and employee motivation. First, the cross-sectional design cannot account for simultaneity effects (i.e., reverse causation) and it is perfectly possible that leaders display different levels of authenticity depending on which employee they are dealing with. Equally, employees who are “more creative” might have higher levels of motivation. Second, the use of questionnaires to measure all variables increases the likelihood of common method bias especially in the case of authenticity and motivation which are both rated by the follower (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). In addition, employee ratings of leader behavior are often influenced by external factors, such as, employee personality, motivated reasoning, organizational culture, and so on (e.g., Hansbrough, Lord, & Schyns, 2015; Lord, Binning, Rush, &
Thomas, 1978). Many of these external factors might also play a causal role in employee motivation and creativity – perhaps extravedted and open employees rate leaders more favourably and are more creative. This brings us to the third class of endogeneity biases: omitted variables. Frequently, studies include just a few variables, omitting many potentially important confounding variables that might influence the nature of the causal effects of leader authenticity on motivation and motivation on creativity.

The consequences of such endogeneity biases can be substantial (Antonakis et al., 2014) and potentially render results uninterpretable, as it is impossible to know whether and to what degree the estimate of the authenticity-motivation and motivation-creativity pathways represent the theorized relationship (i.e., the causal effect) or the unanticipated relationship (i.e., endogeneity biases). As a result, the estimate obtained may be overestimated, underestimated, the opposite sign (i.e., positive instead of negative), or even the opposite direction (i.e., the ‘outcome’ causes the ‘predictor’). In short, the typical leadership-creativity/innovation study is likely to produce biased estimates of causal effects. We need to improve if we are to produce meaningful theory and accurate policy recommendations. There are two well-established study designs that can combat endogeneity biases and provide meaningful estimates of causal relationships: experimental designs and the use of instrumental variables (Antonakis et al., 2010, 2014; Fischer et al., 2017).

Experimental designs

Randomized experiments are the gold standard method for estimating causal effects (e.g., Antonakis et al., 2014). By randomly drawing participants from the population and randomly assigning them to different experimental groups, it becomes highly likely that participants across the different groups will be matched on most characteristics. Thus, when delivering an experimental manipulation to one group but not the other, the researcher can be confident that differences in performance between the groups are due to the manipulation and only the manipulation. Despite the fact that randomized experiments offer the most secure method of estimating the causal effects so central to research concerning leadership, creativity and innovation, only seven studies in our sample used experimental designs (Boies et al., 2015; Chen et al., 2011; Herrmann & Felfe, 2013; Jaussi & Dionne, 2003; Sosik, Kahai, & Avolio, 1998; Sosik, Kahai, & Avolio, 1999; Visser, van Knippenberg, van Kleef, & Wisse, 2013).

There are two types of experimental design used within these studies. The first compared two experimental conditions. For example, Herrmann and Felfe (2014) and Sosik et al. (1999) assigned one group of student participants to a transactional leader and the other group to a transformational leader. Herrmann and Felfe (2014) found that transformational leadership elicited higher levels of creativity than transactional leadership. Interestingly, Sosik et al. (1999) found that transactional leadership led to greater levels of creativity through increased flow, whereas transformational leadership did not. The second design is similar but includes a control group. For example, Boies et al. (2015) assigned participants, working in teams, to one of three experimental conditions with a leader that exhibited inspirational motivation, intellectual stimulation, or an impersonal tone and neutral facial expression (i.e., a control condition). They found that teams working under leaders who exhibited inspirational motivation or intellectual stimulation exhibited significantly higher levels of creative performance than those in the control condition. They also found that levels of creative performance were significantly higher for teams working under the leader who exhibited intellectual stimulation than those working under the leader who exhibited inspirational motivation. Promisingly, all seven of the experimental studies demonstrated causal effects of leadership upon creativity or innovation.

Given the ability of experimental designs to estimate causal effects between variables, we strongly advocate further studies such as those discussed above. However, when designing experiments researchers must pay attention to addressing two key issues: estimating accurate experimental effects through the use of fair comparisons and addressing concerns of ecological validity.

The problem of unfair comparisons refers to designs in which an experimental treatment is compared to a passive control group (Cooper & Richardson, 1986). In such instances, the treatment group can exhibit significant effects due to placebo or expectancy effects. Instead, researchers should ensure that they compare any leadership intervention with a relevant and active comparison condition, which controls for unintended influences on results across treatment groups and provides an estimate of the relative effects of a treatment condition compared with a competing approach (e.g., Chambless & Hollon, 1998). Thus, we recommend the use of randomized experimental designs with multiple treatment conditions and an active control condition (e.g., Boies et al., 2015).

A second important design issue pertains to a longstanding debate regarding the concept of ecological validity, with skeptics arguing that experimental designs do not realistically simulate organizational settings (see Hauser, Linos, & Rogers, 2017). Indeed, experiments published in the organizational literature have often been criticized for using student samples, unrealistic tasks, and failing to reflect realistic leader-follower interactions (e.g., Baumeister, Vohs, & Funder, 2007). Such criticisms are often legitimate and, thus, we suggest three design elements that can help to mitigate these concerns and produce meaningful results.

First, experiments should use realistic and consequential tasks that simulate the need for creativity and innovation as required within organizational settings. For example, many divergent thinking tasks, used to assess the quantity and quality of creative ideas, are completely unrelated to organizational endeavours. For instance, Visser et al. (2013) asked their participants to write down as many different possible uses for a glass of water. Future research should seek to develop protocols for divergent thinking tests that use realistic scenarios (e.g., staff shortages, market competition, financial underperformance). Such research might also and assess participants’ ability across the main stages of creative problem solving (i.e., problem identification, idea generation, idea selection, and implementation planning).

Second, participant incentivization may increase the external validity of experiments (e.g., Hertwig & Ortmann, 2001). The experimental studies in our sample used designs with non-consequential tasks in low-stakes scenarios. For instance, Jaussi and Dionne (2003) asked student participants to develop and present arguments related to ‘the abolition of grades in undergraduate education’. Although participants might have found the task interesting, it is hard to argue that the stakes for performance were high or that participants were motivated as they would be in a real workplace. Incentives, including inducing competitiveness, certificates of completion, providing performance feedback, team-member approval ratings, and financial payments have all been used to increase the ecological validity of experiments (e.g., Lönnqvist, Verkasalo, & Walkowitz, 2011). Monetary incentives are particularly popular because they ensure participants are taking decisions with real economic consequences and thus increase the likelihood that participants perceive the task as consequential and experience realistic emotional reactions (e.g., Falk & Heckman, 2009). However, researchers must carefully consider the use of financial incentives when examining creativity because they induce extrinsic motivation (see Section 3: mediating mechanisms), which might interfere with the experimental effects of interest.

Third, one can move beyond the laboratory and use quasi-experimental designs such as field experiments (Hauser et al., 2017). Field experiments are based within organizations, use real employees and thus can estimate experimental effects within real settings, using high stakes tasks while accounting for complex relationships (e.g., long-standing relationships with leaders) that are difficult to simulate in laboratory settings (e.g., Ilanetz & Staats, 2016). None of the
experimental studies in our sample were field experiments and all used student samples. However, field experiments have been used successfully within leadership research. For example, Dvir, Eden, Avolio, and Shamir (2002) conducted a longitudinal, randomized field experiment in which leaders trained in transformational leadership (experimental condition) had greater impact on followers' development and performance than did leaders trained in eclectic leadership (active control). Despite the aforementioned advantages of field experiments, there are also practical drawbacks and numerous threats to their internal validity. The ecological validity offered by field experiments comes with a loss of experimental control relative to laboratory experiments (e.g., difficult to ensure true randomization and blind conditions). Nevertheless, field experiments are underutilized and provide much stronger tests of causal effect than do the survey-based designs that are typical within the leadership, creativity, and innovation literature.

**Instrumental variables**

One can deal with the issues of endogeneity within cross-sectional and longitudinal field studies by the use of instrumental variables (Antonakis et al., 2010). Within our sample, not a single study utilized instrumental variables, suggesting a lack of awareness within organizational research. Instrumental variables are exogenous predictors (i.e., variables that influence but are not influenced by the model) of an endogenous predictor (i.e., a predictor that relates to an outcome as theorized but also in some unanticipated way, e.g., common method bias). Recall our authenticity → motivation → creativity example, in which leader authenticity and employee motivation were endogenous. In this example, we could use instrumental variables to separate out the endogenous component of leader authenticity and employee motivation. Then using an appropriate model (e.g., structural equation model, 2SLS) we can essentially remove the endogenous association (i.e., that due to common method bias or omitted causes) between authenticity and motivation, meaning that the causal effects can be estimated accurately (see Antonakis et al., 2010, 2014, for technical details).

Although the use of instrumental variables is relatively straightforward, finding appropriate instrumental variables is less so (Larcker & Rusticus, 2010). Instrumental variables must strongly predict the endogenous predictor and must only be related to the outcome variable through their effect on the endogenous predictor. These criteria rule out many established organizational variables such as cultural or organizational structure variables or perhaps even economic conditions (i.e., the presence or absence of recession) because all are likely to influence both leader behavior and employee creativity or innovation; leaving a relatively short list of instrumental variables from which to choose. Antonakis et al. (2010) provide some example instruments, including individual differences that have a substantial genetic component (e.g., cognitive ability, personality), demographic or biological factors (e.g., age, sex, height, hormones), or geographic factors. Some of these variables will likely be ‘stronger’ instruments than others (i.e., more exogenous; Larcker & Rusticus, 2010). For example, although self-ratings of personality traits are moderately heritable (40–55% range; Bouchard & McGue, 2003), personality expression varies across contexts (Fleeson & Jayawickreme, 2015) and trait levels of personality change over time (Roberts, Walton, & Viechtbauer, 2006). On the other hand, cognitive ability is highly stable and heavily heritable (Bouchard & McGue, 2003). Thus, although both are useful instruments, cognitive ability can be considered a stronger instrument than personality (Larcker & Rusticus, 2010). However, regardless of the strength of the instrument, within psychological endeavours none are likely to be very strong predictors of leader behavior and so one would probably need to measure two, three, or more. Nevertheless, the trade-off in survey length is well worth the increased empirical accuracy. Along with Antonakis et al. (2010, 2014) and Fischer et al. (2017), we strongly advocate the use of instrumental variables when examining causal process models within cross-sectional or longitudinal field studies.

**Section 5: measuring creativity and innovation**

“...the primary issue to hamper creativity research centers around the lack of a clear and widely accepted definition for creativity, which, in turn, has impeded efforts to measure the construct.” (Batey, 2012, p. 55)

Theory and measurement are the core aspects of any science, with the development of accurate, precise and (study-) appropriate measures the fundamental base for all other empirical endeavours (Hughes, 2018). Unfortunately, as the quote atop this section notes and as we discussed in Section 1, defining and measuring creativity and innovation has proven a genuine challenge for researchers (Anderson et al., 2014; Batey, 2012). Numerous reviews of creativity and/or innovation at work have made comment regarding measurement, documenting popular measures (e.g., Anderson et al., 2014), commenting upon trends regarding the source of ratings (e.g., self-ratings or supervisor-ratings; Harari, Reaves, & Viswesvaran, 2016; Ng & Feldman, 2012), or discussing specific measurement-based issues (e.g., common-method bias; Zhou & Shalley, 2003). However, all have stopped short of critical reviews of the measures themselves, which we believe to be an important oversight. Such a review would aid researchers in choosing appropriate measures for their study, potentially shed some light on the conflicting findings within the literature, and provide guidance regarding future measure development. Accordingly, we too document the nature of the measures used within our article sample but also examine frequently used measures with regard to what is commonly termed ‘validity’.

**Which measures are used?**

Studies that examine links between leadership and creativity/innovation have employed a diverse range of measures such as self-rated psychometric scales, other-rated (i.e., colleague or supervisor) psychometric scales, counts of objective criteria, and experimental measures. Despite the overall diversity, the preponderance of studies utilized psychometric questionnaires of some sort and so these measures deserve special attention, and we will return to them shortly. First, however, we consider non-survey based measures.

Only ten studies assessed creativity and innovation through non-survey based measures. Five of the experimental studies identified used non-survey-based measures of creativity and innovation, typically some variant of a divergent thinking test. In essence, divergent thinking tests require participants to generate multiple alternative answers/suggestions/solutions to open-ended problems, and in doing so, they assess the central component of creativity: the generation of ideas (Mumford, Marks, Connelly, Zaccaro, & Johnson, 1998). For example, Sosik et al. (1998) examined the effects of transformational leadership upon creativity exhibited using an online brainstorming tool. Leadership was manipulated by having confederates behave in a manner that was consistent with either high or low transformational leadership. Creativity was assessed through judge ratings of different aspects of the ideas generated during the brainstorming. Specifically, judges assessed idea fluency (i.e., total number or original ideas), flexibility (i.e., range of ‘categories’ of ideas), originality (i.e., novelty of ideas), and elaboration (i.e., suggestions to improve initial ideas). They found that transformational leadership was associated with increased originality and elaboration. Three other studies used variants of divergent thinking tasks (Herrmann & Felle, 2014; Jaussi & Dionne, 2003; Visser et al., 2013) and one used a building block task (Boies et al., 2015).

We noted previously how useful experimental designs are and divergent thinking tasks provide an appropriate and well-established method for assessing creativity in experiments (Batey, 2012). Divergent thinking tasks can be scored objectively (e.g., fluency: counts of all ideas, originality: counts of unique ideas) or subjectively (e.g., expert-ratings of originality or quality), and there is much debate regarding
which approach is best (cf. Amabile, 1996; Plucker, Qian, & Schmalingsee, 2014; Silvia et al., 2008). Currently, best practice guidelines would suggest some combination or product (e.g., fluency/originality) of both, with all methods being useful to varying degrees (Plucker et al., 2014). Which is most appropriate will be context and study dependent. For instance, if the organizational model requires a large quantity of proposed ideas (e.g., fashion design) then fluency might be useful, if the model requires highly original ideas more than pragmatic ones (e.g., marketing) then originality might be most useful (Plucker et al., 2014; Runco, Abdulla, Paek, Al-Jasim, & Alsuwaidi, 2016). Thus researchers should identify and justify which is the most appropriate approach for their study and task, and as discussed in Section 4: study design, should obtain divergent thinking scores using realistic tasks within consequential environments.

In addition to divergent thinking tests, six studies utilized organization-specific markers of creative or innovative performance. Examples include archival records of employee ideas, published research reports, product innovation sales as a proportion of total sales, ratio of product innovation sales to product innovation development costs, and paid ‘creativity bonuses’. The use of such non-survey data is generally possible and provides tangible ‘real-world’ assessments of organizational performance. However, it is important to note that such metrics typically do not provide insight into the processes and mechanisms that facilitate creative or innovative performance. It is also important that researchers use company data effectively and in-line with theory. For example, Jung, Chow, and Wu (2003) and Jung, Wu, and Chow (2008) used various composite scores calculated from the number of patents, research and development expenditure, and ratings of organizational innovation. In essence, this approach mixed objective and subjective metrics as well as inputs (research and development expenditure) and outputs (patents). Scoring variables in such an atheoretical manner is unwise and likely hides important nuanced relationships that are apparent when different variables are examined independently (e.g., Tierney, Farmer, & Graen, 1999).

When objective data are not available and experimental designs are not appropriate, the most flexible approach is the use of psychometric scales. The most commonly used creativity scales purport to assess ‘creativity’ (Zhou & George, 2001; 37% of studies), ‘employee creativity’ (Tierney et al., 1999, 17% of studies), and ‘creative performance’ (Oldham & Cummings, 1996, 7% of studies). The most commonly used innovation measures purport to assess ‘innovative behavior’ (Scott & Bruce, 1994, 11% of studies) and ‘innovative work behavior’ (Janssen, 2000, 16% of studies; De Jong & Den Hartog, 2010, 2%). In the next section, we examine each of these measures in more detail.

It is also common practice to modify items slightly, take a subset of one scale’s items, or combine items from multiple scales. Usually, the rationale is that the items do not fit the context or that the need for brevity is great. Changing items so that they better fit a particular context is not necessarily problematic; indeed, it is preferable to using inappropriate items. However, a modified measure is not synonymous with the original measure and researchers should make efforts to examine the psychometric properties of the new scale and check that it is actually measuring what they hope it is measuring. However, not a single study that used a modified scale conducted any thorough evaluations. Similarly, taking a sub-sample of items is not sufficiently problematic. In fact, if the construct is unidimensional and the item subset retains reasonable coverage of the construct, each item is an equally reliable indicator (i.e., has roughly equal factor loadings), and the scale provides similar psychometric properties (e.g., factor structure, reliability), then shortened scales can be very useful indeed (e.g., Tokarev, Phillips, Hughes, & Irwing, 2017). However, creativity and innovation are multidimensional constructs and the approach to item selection often appears piecemeal, if it is discussed at all.

Are workplace creativity and innovation scales accurate and appropriate?

Typically, scale evaluations focus on the concept of ‘validity’. However, the word ‘validity’ is widely used but rarely well-defined (Newton & Shaw, 2016). As a result, many researchers have come to regard validity as a complicated issue – which it is not – and clear treatises on validation practices are hard to find (Borsboom, Mellenbergh, & Van Heerden, 2004). In the current paper, we are guided by Hughes’ (2018) Accuracy and Appropriateness model of test evaluation, which contends that validation is an on-going process of scale evaluation that consists of two sequential lines of enquiry: first, establishing whether a test accurately measures what it purports to measure; second, establishing whether the use of a test for a given purpose is appropriate. In the following sections, we review the most commonly used creativity and innovation scales with regard to their accuracy and appropriateness.

Accuracy: construct representation

The accuracy of a measure is established when item content provides representative coverage of the theoretical construct, responding to the items elicits the desired participant responses, the structure of the scale matches the theoretical structure (i.e., factor structure), the measure functions equivalently across groups, and it demonstrates convergent relationships with scales assessing the same construct. Evidencing content representativeness or content accuracy, is the first step, and involves demonstrating a match between the theorized content of the construct (i.e., construct definition) and the actual content (i.e., items) of the psychometric scale (Nunnally & Bernstein, 1994). So, what do the most commonly used scales, actually assess? To address this question, five subject matter experts1 (SMEs) conducted an item-level review of six scales, three that ostensibly assess creativity and three that ostensibly assess innovation. Each of the SMEs independently coded the items with respect to two different concerns.

First, SMEs rated whether each item assessed creativity, innovation, both, or neither in accordance with the integrative definitions generated in Section 1: defining creativity and innovation. Specifically, an item was considered to assess ‘creativity’ if it referred to the generation of novel ideas, and to assess ‘innovation’ if it referred to processes germane to the implementation of ideas. If an item assessed some combination of these, it was rated as a measure of both creativity and innovation, and if the items assessed elements outwith idea generation and implementation, it was considered to assess neither creativity nor innovation.

Second, guided by Batey’s (2012) measurement framework, SMEs examined which element of creativity or innovation was assessed. Batey’s framework acknowledges that creativity and innovation are not singular static constructs; rather they are the product of a process undertaken by a person or persons within an environment (press). Accordingly, measures of creativity and innovation can assess one (or more) of four facets:

- **Person/Trait**: A person’s characteristics or traits that are conducive to creativity or innovation.
- **Process**: Behaviors, actions, and cognitive processes that a person/team engages in when attempting to generate and implement creative ideas.
- **Press/Environment**: The features of the environment within which creativity/innovation takes place.
- **Product**: The creative ideas generated or innovative outputs implemented.

1 Including the first author, two creativity and innovation researchers who have published extensively, and two PhD students researching the conceptualisation and measurement of innovative work behavior and team creativity.
Once completed, we summed the ratings across the five SMEs and coded items according to the predominant view. For every item examined, at least four of the five SMEs agreed. The summary of this analysis is contained in Table 5, with example items displayed in Table 6.

With regard to Table 5, the most important observation pertains to the fact that there is considerable variation both within and between creativity and innovation scales. All six scales contain items that assess creativity, innovation, or a mixture of both. In addition, three of the scales contain items that assess neither creativity nor innovation. For the creativity scales, 48% of items assess some element of creativity, 20% assess innovation, 16% assess both creativity and innovation, and a further 16% assess neither creativity nor innovation. A similar picture pertains to the innovation scales, with 64% of items assessing innovation, 16% assessing creativity, 16% assessing both, and 4% assessing neither. Put simply, scales that ostensibly assess creativity also assess innovation and vice versa. Such overlap might not necessarily be a problem if it was explicitly acknowledged, scales were labelled ‘creative and innovative behaviour’, and the items were not treated as unidimensional and sum-scored. However, none of those things are the case in current practice, rather the scales reflect problemmatic levels of conceptual confusion, which lead to two further fundamental concerns.

First, the two bodies of literature – creativity and innovation – are considered related but distinct and recent reviews have called for further integration (Anderson et al., 2014). However, our analysis reveals that the two fields are already synonymous. If a researcher has used the popular Zhou and George (2001) creativity scale and summed the items to form a single scale score, then that score is approximately 40% creativity, 30% innovation, and 25% irrelevant content (see Table 5). So what exactly is that total score: creativity, innovation, or both? All of the scales analysed appear to offer a very broad, non-specific measure of various elements of creativity and innovation at varying ratios. Thus, it seems the most sensible conclusion is ‘both’. Indeed, the item content analysis explains why it has proven difficult to separate the two constructs empirically. For instance, a recent meta-analysis by Harari et al. (2016) combined measures of creativity and innovation into a single category, namely, creative and innovative performance, and remarked that even when the two were separated there was virtually no difference in the pattern of correlations observed. Given the analysis of the item content, this similarity is unsurprising, and considering the two sets of scales as markers of a broader creativity and innovation variable, as Harari et al. (2016) did, seems appropriate.

Harari et al. (2016) state further that this single variable represents a job performance domain, noting specifically that it does not refer to the processes that lead to performance. This brings us to the second point of particular note. With the exception of Oldham and Cummings (1996), there is variation in the extent to which scales assess judgments of persons/traits, processes/behaviors, and products/performances. Overall, 16% of items within ostensible creativity scales assess the person, 24% processes, and 60% products. For ostensible innovation scales, 4% assess persons, 68% processes, and 28% products. Thus, the premise of Harari et al.’s (2016) classification is flawed. These scales do not simply measure creative or innovative performance but also the processes that lead to performance and some personal characteristics associated with creativity and innovation. In fact, when looking within rather than across scales, around 50% of the items within scales labelled ‘creativity’ and between 16% and 30% of items within scales labelled ‘innovation’ assess products or performance. Yet again, it is not entirely clear what these scales represent, especially when sum-scored.

Interestingly, within each measure of innovation, the proportion of...
process-related items is greater than the proportion of product-related items (with the product items usually being those that assess creativity not innovation), and the opposite is true of creativity scales. The field would benefit from measures that offer nuanced assessment of creative and innovative processes and creative and innovative performance/products. The latter are crucial for assessing overall competence and the former are crucial if we wish to assess how employees create/innovate in order to build meaningful training programmes and interventions.

In sum, psychometric scales of workplace creativity and innovation mix creativity and innovation items along with person, process, and product items. Thus, current measures are simply not strong enough and we need to develop new ones. The lack of clarity within these measures is likely reflective of problematic definitions that confused creativity and innovation, and used antecedents and outcomes to define them both (see Section 1: defining creativity and innovation). The blurring of person, process and product items is problematic, because these scales are often used as outcome variables but contain content that directly overlaps with predictor variables (e.g., personality). In addition, mixing creativity and innovation items that tap each of the 4Ps promotes contradictory findings within the literature (e.g., Harari et al., 2016). Indeed, previous studies (not conducted within the leadership field), have shown that different aspects of creativity and innovation have different antecedents and differential relationships with other variables (e.g., Axtell et al., 2000; Magadley & Birdi, 2012; Rank et al., 2004). Thus, definitional confusion has led to inaccurate and imprecise measurement that has limited meaningful theoretical discoveries and advances.

Accuracy: scale development and psychometric properties

In addition to item content that accurately reflects the theoretical construct, scale accuracy can also be demonstrated through a range of commonly utilized psychometric analyses, such as factor and invariance analyses (see Hughes, 2018). Unfortunately, such analyses were not commonly performed during the development of the six scales in Table 5. With the exception of De Jong and Den Hartog (2010), all scales were developed in small (N < 200), homogenous samples, collected from one or two organizations, and none were subject to any structural analyses. In contrast, De Jong and Den Hartog (2010) developed their scale within two separate samples (n = 81 and n = 703) and examined both exploratory and confirmatory factor models. However, the overall picture is one of poor scale construction: we can and must do better than using scales with mixed item content developed without the application of rigorous psychometric analyses (see Hughes, 2018, and Irwing & Hughes, 2018, for holistic guides to scale development).

The current review makes very clear that we are in need of new scales to assess workplace creativity and innovation: new scales that offer clear facet-level measurement and scales that distinguish between person, process, and product. In particular, given the overarching goal of this field is to build models that explain how leader behaviors can facilitate/hinder creativity/innovation, we need new measures that include behavioral items that describe the activities that employees/teams/organizations engage in to generate and implement creative ideas. Without such scales, it will be difficult to disentangle the undoubtedly complex set of relationships between leadership and different elements of the creative and innovative process.

Appropriateness: are workplace creativity and innovation measures appropriate for future research?

Only once a scale has been shown to accurately capture its intended target can it be considered for use in theory testing/building or decision-making. Establishing whether or not a measure is appropriate is context- and goal-specific, but usually involves assessment of the scale’s relationships with other variables (e.g., predictive properties), the feasibility of scale use (e.g., length, cost), and the potential consequences of scale use (Hughes, 2018). The simple statement here, given the nature of our review, is that none is particularly appropriate for future research. However, this position would be somewhat overzealous, and indeed we feel that use of some of these scales can be justified in the right circumstances. Thus we make some tentative recommendations regarding which of these six measures we would use.

If one wants to assess employee performance in both creativity and innovation equally, at the broadest possible level of abstraction, then perhaps the best option is the 3-item scale by Oldham and Cummings (1996). This scale will be especially useful when constraints on survey length are particularly stringent. However, because Oldham & Cummings' items are ‘double-barrelled’ (i.e., ask two things per item) it is likely that ratings will contain a non-negligible proportion of measurement error that hides differences between creative and innovative employees (Irwing & Hughes, 2018). If one wants to assess employee creative processes and performance (or products) then the 9-item scale by Tierney et al. (1999) would appear to be the most appropriate. If one wants to assess a combination of creative performance and innovative behavior/processes then De Jong and Den Hartog's (2010) measure looks most promising. In this case, we would urge researchers to analyse the scales’ four sub-factors (opportunity exploration, idea generation, idea championing, idea application) separately, as well as analysing a single latent factor of ‘global innovative work behaviour’, loaded by these four factors (see De Jong & Den Hartog, 2010, Fig. 1).

Despite the above recommendations, our clear and unequivocal call, is for the relatively urgent (but thorough) development of new, theoretically salient and psychometrically robust measures of workplace creativity and innovation. Given the development of new scales will take some time, we advocate the use of existing scales as outlined above but urge researchers to exercise vigilance and explicitly acknowledge and discuss the implications of the limitations of the measures they use.

Discussion

Creativity and Innovation are vital for organizational success and are intriguing topics to research. Leadership is considered to be a major contextual factor that influences employee creativity and innovation (e.g., Anderson et al., 2014; Shalley & Gilson, 2004; Tierney, 2008), and research in this area is burgeoning, with 85% of the studies included in our review published in the last 10 years. The growth of leadership-creativity/innovation research has been swift and largely exploratory, with individual studies typically not building systematically toward a unified body of evidence. Throughout our review, we have identified major trends, provided broad frameworks and taxonomies to give some structure to the literature, and highlighted how future research can move this important area of research forward in a systematic and rigorous manner.

Below, we propose specific calls for future research in five main areas that are crucial to making progress. We organize these calls according to the major sections of our review. However, we present them in reverse order. We begin with measurement. Without accurate and appropriate measures of creativity and innovation, all other empirical endeavours are futile. Next, we consider study design. Without appropriate study designs, researchers cannot be confident in testing the causal claims that are so central to this field. Next, we consider leader variables, moderators, and last we consider mediating mechanisms. We hope that this order of presentation makes explicit that the most urgent need for future research concerns the development of nuanced measures that accurately capture the different elements of creativity and innovation and the careful consideration of study design. Without these foundational qualities, the scientific merit and practical utility of future studies will be limited.
Measuring creativity and innovation

Across psychological and organizational research, we are sometimes so eager to conduct theoretical research that we play ‘fast and loose’ with construct definitions and the procedures we follow when translating these definitions into measurement scales. Indeed, some of the most widely used creativity and innovation measures were developed as a minor part of a field study, rather than as a standalone and thorough empirical endeavour (e.g., Scott & Bruce, 1994). As a result, we are doing ourselves a disservice, wasting time, money, and other resources, because, without high quality measurement, all other empirical efforts are conducted in vain. Our review of extant measures provides a clear message: we need new tools to assess workplace creativity and innovation. Below are our specific calls for future research in this domain:

1. New psychometric scales that:
   a. Assess the key stages of creativity and innovation in enough detail to provide fully construct representative measurement.
   b. Assess the different facets of creativity/innovation (i.e., person, process, & product). For example, one could measure the process of idea promotion or one could assess the quality (i.e., product/performance aspect) of idea promotion. In particular, there is a need for behavioral/process measures that assess how individuals or teams generate and implement novel ideas.

2. Once we have new scales that measure creativity and innovation exclusively, research can begin to examine how the two interrelate within the workplace.
3. Once we have new measures, research can begin to assess the common and unique antecedents of creativity, innovation and their sub-processes. One important antecedent will, of course, be leader behavior.
4. There is a need to develop realistic divergent thinking tasks that allow for the assessment of different aspects of creativity and innovation.

Study design

Every study within our sample was concerned with assessing casual effects but most employed study designs poorly suited to estimating causal models due to their susceptibility to endogeneity biases (Antonakis et al., 2010, 2014; Fischer et al., 2017). In order to improve the rigour of future research, and thus, build more reliable and useful theories and practical recommendations, we make the following suggestions for study design:

5. It is imperative that researchers establish causality using designs that are capable of doing so. The best option here is the experimental design.
6. Where experimental designs are not possible or inappropriate, field studies need to do everything possible to identify instrumental variables and use them to control for endogenous variance with their models.
7. Move away from the reliance on cross-sectional designs instead using longitudinal designs, with theoretically appropriate time-lags (Fischer et al., 2017).
8. It seems to us that the best possible design is a multi-study paper including two or more studies. First, a randomized controlled experiment that establishes the nature of causality for each part of the model. Second, a cross-sectional/longitudinal field experiment or survey-study, that, where necessary uses instrumental variables and appropriate time-lags. Such studies will provide a check on the ecological validity of the observed experimental effect.

Leadership: main effects

Our review highlighted that many of the key conceptual challenges associated with the wider leadership literature are present in the leadership, creativity and innovation literature. Specifically, numerous ‘positive’ leader approaches correlated positively and ‘negative’ leader approaches correlated negatively with creativity and innovation. However, it is unclear which leadership approaches are the strongest predictors because the literature has largely failed to examine the relative contribution of different leadership variables. The recent proliferation of “positive” forms of leadership (i.e., servant, authentic, empowering, and ethical leadership) has served to exacerbate the problem. The uniform pattern of correlations, regardless of the nature of the leadership style, suggests that we might be measuring overall attitudes toward leaders rather than actual behaviors (Baumeister et al., 2007; Lee, Martin, Thomas, Guillaume, & Maio, 2015). Future research needs to address the lack of empirical clarity by focusing on the distinctive elements of leadership approaches and their relative and incremental effects. In addition, research should consider moving away from broad leader “styles” to consider more nuanced behavior, which will increase our understanding of the basic building blocks of leader influence.

Leadership: moderators

15. Try to replicate the moderating effects found in single studies, across leadership variables and contexts.
16. Develop a clear theoretical framework to classify moderating variables that can be used to guide future research in a similar manner as our mediator classification.
17. Greater focus on broader context, for example, the role of industrial context.

Leadership: mediators

Our review identified many mediating variables and a range of solid theoretical rationales to expect that a leader’s influence on creativity and innovation is mediated. Specifically, we used our review to build a theoretically-driven five-class categorization of mediators: motivational, affective, cognitive, identification, and relational. Future research should use these classes to compare mediators both within and between categories to build systematic understanding of how leaders influence both creativity and innovation. Below are further calls for future research:

18. Avoid over-emphasis on motivational processes to the detriment of other understudied mechanisms (e.g., affective, cognitive).
19. Explore the relative effects of intrinsic versus extrinsic motivation for both creativity and innovation. In particular, focus on creativity-contingent rewards and how these relate to intrinsic motivation.
20. Greater efforts to exploring competing mediating pathways, both within category (i.e., self-efficacy versus psychological empowerment) and between categories (i.e., motivational versus cognitive).
21. Explore how leaders develop followers’ cognitive skills/abilities, as opposed to how they leverage them.
22. Examine how different mediators relate to different aspects of creativity and innovation (e.g., idea generation vs. idea implementation).
23. Determine which leadership variables are most predictive of which mediators.

24. Greater focus on team and organizational processes.

Conclusion

Our review has shown that leadership, creativity, and innovation research is an active and growing area of enquiry that has yielded numerous interesting and intriguing findings. In particular, there is clear theoretical and empirical evidence demonstrating that leadership is an important variable that can enhance or hinder workplace creativity and innovation. Thus, further study is warranted to build a more precise understanding of which leader behaviors are most important and to identify the mechanisms through which these leader behaviors carry their influence. However, we call not for more of the same, not for more small exploratory studies, but for rigorous and more comprehensive studies. We call upon researchers to look afresh at things often taken for granted and in doing so, follow the wisdom of Albert Einstein: “To raise new questions, new possibilities, to regard old problems from a new angle, requires creative imagination and marks real advance in science.” Specifically, we urge researchers to think creatively to address the measurement, study design, and theoretical concerns discussed above, so that the field can build and examine theoretical propositions in a manner that produces accurate and reliable policy recommendations.

References


*Strickland, S., & Towler, A. (2011). Correlates of creative behavior: The role of leader-


*Sun, L.-Y., Zhang, Z., Qi, J., & Chen, Z. (2012). Empowerment and creativity: A cross-

*level investigation. The Leadership Quarterly, 23, 55–65.


*Trang, N. N., Nghi, P. T., Soldier, L. L., Hai, T. V., & Kim, W. J. (2014). Leadership, resource and organizational innovation: Findings from state and non-state en-


*Volmer, J., Spark, D., & Niessen, C. (2012). Leader-member exchange (LMX) job au-


*Zhou, H., & Towler, A. (2014). Supervisor–subordinate relationship, differen-


*Zhichao, C., & Cui, L. (2012). Impact of mean LMX on team innovation: An empirical investiga-


*Zhuo, Q., & Pan, W. (2015). A cross-level examination of the process linking transfor-