

FOCAL ARTICLE

An epistemology for assessment and development: How do we know what we know?

Deborah E. Rupp¹, George C. Thornton III², Tiffany M. Bisbey³, Anna N. Hoover¹, Eduardo Salas⁴, and Kevin R. Murphy⁵ 

¹George Mason University, Fairfax, VA, USA, ²Colorado State University, Fort Collins, CO, USA, ³The George Washington University, Washington, USA, DC, ⁴Rice University, Houston, TX, USA and ⁵University of Limerick, Limerick, Ireland

Corresponding author: Deborah E. Rupp; Email: Drupp2@gmu.edu

(Received 20 December 2023; accepted 3 January 2024; first published online 18 March 2024)

Abstract

To make informed decisions, assessment theorists, researchers, and practitioners can evaluate the overlap among (1) relevant theories, (2) empirical contributions, and (3) best practices. Unfortunately, such a task may seem daunting due to the so-called science-practice gap, which can thwart collaboration among these parties. This paper presents an epistemology for delineating the importance of integrating these three sources of knowledge. We then apply this epistemology to show that our current knowledge of assessment and development topics are well integrated in some places, but still quite lacking in others.

Keywords: Assessment centers; training; science-practice gap; epistemology; assessment; development

Introduction

An abundance of information on effective assessment and development is readily available via a variety of outlets. For example, there exists peer-reviewed articles that describe the latest theory and research, best-selling books, expert reports on industry trends, and social media testimonials of executives with decades of practical experience (Ployhart & Bartunek, 2019). Of course, each of these sources is not equally valuable; and even the most valuable sources have limitations. Thus, it becomes the job of responsible consumers of assessment and development knowledge to sift through the vast array of sources to identify unbiased information applicable to their purpose (Lowman & Cooper, 2018). For practitioners, their purpose may involve selecting appropriate techniques to assess and develop human talent; whereas scholars may theorize about and investigate the techniques practitioners implement.

In doing so, knowledge consumers may struggle to gain *accurate* knowledge about assessment and development, especially when many variations of a practice exist (e.g., among various performance indices, evaluation schedules, and feedback approaches; Murphy et al., 2019). Some practitioners might resort to off-the-shelf products marketed as “scientifically validated” and “evidence-based” with limited access to the supporting evidence. Even in the most robust scientific efforts, academics may lack first-hand experience and understanding of organizational realities, or the feasibility of the recommendations they make; leading to a lack of clarity about what exactly should inform future research and theory (Rynes, 2012; Rynes et al., 2001). Gaining the wealth of knowledge available is ultimately throttled by the lack of accessibility of different sources of information, conflicting information, and insufficient detail to infer generalizability. To aid researchers and practitioners in addressing this challenge, we provide a three-part epistemology aimed at gaining the most complete information about assessment and development practices.

Rather than presenting new information, in this paper we present and support a framework for processing information. Further, we explain how this framework for acquiring and processing information, or epistemology, offers greater clarity or a new perspective on known issues.

Ways to gain knowledge: epistemology

For centuries, philosophers of science have studied the basic question of how one comes to know something. To *know* includes being aware of information and its accuracy, and using multiple processes and sources to gather it. Epistemology, a sub-field of philosophy, describes how knowledge is acquired and justified (Audi, 2011). An epistemology can range from accepting doctrines from authority, to conducting empirical research, to gaining insight from personal experience. In this paper, we apply an epistemology to assessment and development to demonstrate its value in gathering the most complete knowledge on various practices within these areas.

Audi (2011) explains that one aspect of epistemology is examining the *justification* of a belief, asserting that we are justified if we have some basis for the belief. There are several bases for believing we know something. Our proposed epistemology consists of three bases for knowledge: theory, empirical research, and observations of practice. *Theory* is a systematically organized set of knowledge, including assumptions, principles, and relationships among concepts (Sutton & Staw, 1995). *Empirical research* involves methods of gathering and interpreting data to uncover or confirm knowledge (APA, 2015). *Practice* allows for knowledge to amass from one's own or awareness of others' direct observations of effective and ineffective techniques (Rupp & Beal, 2007). These varying ways in which we gain knowledge differ by provided certainty about verified facts, connections with other ways of knowing, and reliance on internal states of awareness or information external to the knowledge seeker (Shieber, 2019).

Our epistemology relies on a combination of more certain pieces of interconnected information derived from coherent theoretical propositions, rigorously conducted empirical research evidence, and effective organizational practice. It represents a divergence from the traditional scientific assumption that valid knowledge is gained only through systematic, deductive research (McLelland, 2006). It is also more amenable to the assessment and development fields, given their inherent blend of science and practice (Benjamin & Baker, 2000).

We argue that this approach is important, especially because the field of personnel assessment and decisions has, historically, eschewed insights from practice (Rynes et al., 2001; Rynes & Bartunek, 2017) while simultaneously publishing "theory" that is often little more than intuitive judgement propped up by post-hoc threading together of past theory developed in the same manner (Rupp et al., 2017). We argue that bringing together our identified three sources of knowledge is key for moving the field forward by illuminating what we know about assessment and development, and the strength of that evidence, through the integration of multiple knowledge sources.

A three-part epistemology

Our proposed epistemology (Figure 1) provides several contributions. First, it provides a unique and organized process for identifying what we do and do not know about assessment and development. Evaluating and integrating information from theory, empirical research, and practice provides an encompassing appraisal of the state of knowledge in an area. Thus, we place the greatest emphasis on the intersection of all three sources. Second, it provides a means for evaluating assessment and development practices based on converging evidence. Finally, our approach calls for investigating relationships between assessment/development practices and their role in the larger contextual system, which research has been criticized of ignoring (Jackson & Shuler, 1995; Johns, 1993; Parrigon et al., 2017).

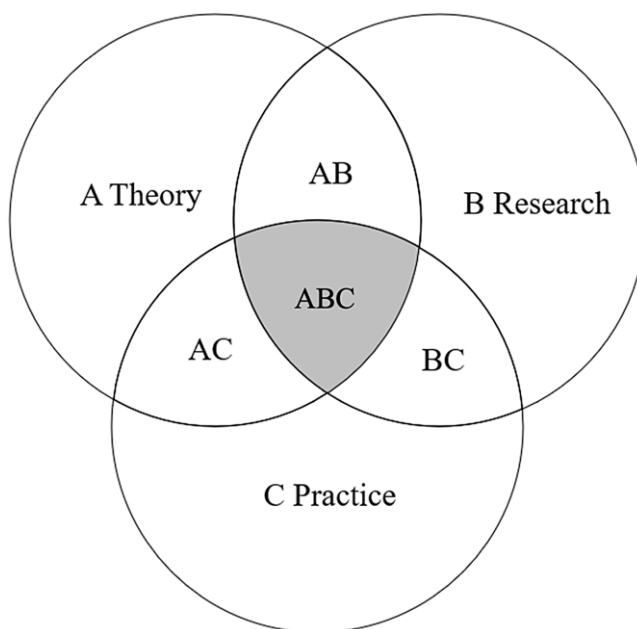


Figure 1. Venn Diagram of the Proposed Epistemology.

In the following sections, we describe (a) the three focal segments of our epistemology (see Figure 1); (b) how knowledge is acquired for each segment; (c) each segment's strengths and limitations; and (d) how each segment contributes to knowledge. Then, we examine what we can gain from a lack of agreement or even conflict between and within each knowledge base. Finally, we showcase convergence among ways of knowing about assessment (specifically assessment centers) and development (specifically training).

Theory

Theory is a set of reasoned beliefs about relationships among variables. Bacharach (1989) states “a theory is a statement of relations among concepts within a set of boundary assumptions and constraints” (p. 496). Theories can impel action and force us to go further than opinions to substantiate our beliefs by asking how we formed ideas, how things work, and how they might be done differently (Nealon & Giroux, 2012).

Advantages and disadvantages of theory

A theory can summarize knowledge accumulated over time across a body of research (Suddaby, 2014), and succinctly capture results of systematic interventions in practice (Locke, 2007). However, relying on theory alone can lead to a neglect of experimentation and observation (Cucina et al., 2014). At times, hypothesized statements may be nothing more than “received doctrine” or “academic intuition.” Furthermore, a theory can bias and impede progress if it continues to be relied upon without substantiation (Hambrick, 2007). A theory is strengthened if supported by research and field observations of fruitful or ineffective practice.

Asking specific questions about a theory can assist in determining its strength: How widely adopted is the theory? Are there other competing perspectives? How significant are areas of disagreement? Has the theory been tested through research and applied in practice? Strong theories have been empirically tested multiple times with replicated findings (Cucina et al., 2014;

Hambrick, 2007; Woo et al., 2017). Untested theories risk devolving into pseudoscientific theories, or “explanations based on conjecture, personal opinion, and limited findings that cannot be called true theories” (Woo et al., 2017, p. 257). Thus, the strength of a theory partly depends on whether empirical research supports the phenomenon of interest.

Empirical research

Empirical research involves gathering data to confirm theoretical relationships or identify new relationships. An emphasis on empirical research as a basis of knowledge is compatible with the recent emphasis on “evidence-based management” (EBM; Barends et al., 2014; Barends & Rousseau, 2018), which refers to a variety of methods on which to base business decisions via empirically gathered information. EBM can assist in raising and deriving the issue of what constitutes evidence, the strength of evidence, and how academics and practitioners can improve the quality of their evidence (Rousseau & Gunia, 2016).

Approaches to empirical research

Deductive research (i.e., the basis of empirical theory testing) is useful when relevant theories exist. However, both inductive and abductive research also fulfill important roles in identifying and explaining phenomena (Locke, 2007). Woo et al. (2017) note that inductive research facilitates exploration of emerging questions that may not yet have theory to support them. Similarly, abductive research makes observations based on data, but also seeks to provide an explanation for the observations (Folger & Stein, 2017). Abductive reasoning typically begins with an incomplete set of observations and proceeds to the likeliest possible explanation for the set. Although organizational research has historically prioritized theory-driven deductive research, inductive and abductive research valuably contribute to the advancement of scientific knowledge through the generation and explanation of new research questions.

Complexifying empirical research

The inclusion of multiple variables can increase the strength of empirical research by modeling more complex and realistic relationships between them (Berry & Sanders, 2018). Also, research can be simultaneously conducted at multiple levels of analysis that are hierarchically nested within each other, such as employees within teams, to better account for the complex reality of relationships in practice (Klein & Kozłowski, 2000; Zhou et al., 2019). Studies with nested data that do not account for multiple levels of analysis may misinterpret results. For instance, there may appear to be no relationship between training and employee performance until accounting for the effect of team membership, which might direct investigation into team-level variables influencing results (Snijders & Bosker, 2011).

In addition to traditional research methods, recent innovations in big-data methodologies, including artificial intelligence and machine learning, provide new possibilities in prediction (Tonidandel et al., 2016). These methods often rely on “organic data,” or large datasets that emerge from ongoing information collection processes (e.g., HRM information systems; Groves, 2011). Organic data often arise from processes designed to tackle practical problems (McAbee et al., 2017), and typically constitute data that would be impossible or prohibitively difficult to collect via traditional research methods (Woo et al., 2020). With increased access to such data and computational power to analyze larger datasets, more complex predictive models can be tested with statistical rigor. However, many researchers caution against “dustbowl empiricism” or atheoretical approaches often consistent with big-data methods, in that conclusions are driven by data collected under uncontrolled conditions. Further, decision-making models “backed” by big-data algorithms have the potential to demonstrate bias (e.g., in assessment for selection, Dastin, 2018), as well as a potential inability to out-perform traditional methods (Hickman et al., 2019).

Meta-analysis

Another type of research, meta-analysis, if carried out in accordance with professional standards (see APA, 2020), can overcome some of the limitations of individual empirical research studies (Huffcut, 2004). Meta-analyses acknowledge the concept of sampling error in research and treat each individual study as a sample from the population, providing a more precise estimate of the actual relationship between variables, assuming most of the studies used to provide the correlations for the meta-analysis are not themselves highly flawed (LeBreton *et al.*, 2014).

Meta-analysis overcomes problems of relying on the results of a single study to make valuable theoretical contributions by examining the full collected set of evidence available. For example, Schmidt and Hunter (1998; Hunter & Schmidt, 2004) produced guidelines for meta-analytic methods that have served as a theoretical foundation for determining the relative predictive ability of various assessment techniques; and offered one of the first comparative analyses of predictor constructs and methods commonly used in personnel selection. An example of the application of meta-analysis is Ones *et al.* (1993), who used the method to test the generalizability of integrity test validities, finding that they can predict several types of counterproductive work behavior, extending previous theory suggesting integrity as a predictor of solely employee theft.

Importantly though, meta-analyses must be continually updated. Sackett *et al.* (2022), after identifying that Schmidt and Hunter (1998) had applied overcorrections for restriction of range, conducted a new meta-analysis using updated methods and research studies as input. They revealed meta-analytic estimates of predictor criterion-related validity that were substantially smaller than those reported by Schmidt and Hunter. Similarly, Van Iddekinge *et al.* (2012) conducted an update to the Ones *et al.* (1993) meta-analysis and found that the predictive validity of the integrity test was much smaller than previously specified. In sum, meta-analyses contribute to knowledge through their summation of individual empirical research studies; however, they also must be updated over time to reflect knowledge gained through new research.

Advantages and disadvantages of empirical research

The advantages of empirical research involve the ability to infer statistical and practical significance of observed relationships, and the confidence inherent in these conclusions. However, there are limitations to what can be learned through individual empirical research. By necessity, the number of variables examined in any one study is limited, and thus cannot completely reflect the complexity of real organizations. Field research in organizations can be expensive and intrusive, and thus, replication is seldom carried out. In laboratory studies, it can be difficult to operationalize a variable to mimic work-related constructs. Moreover, research samples may not always match the organizational populations to which the research seeks to generalize. A tradeoff exists between the control a particular methodology allows, and the generalizability of the results (Cook & Campbell, 1979). If the research aim is to generalize to a variety of contexts, a strong research study may prioritize generalizability at the expense of control. Alternatively, if the goal of the research is to determine causal relationships, strong research will prioritize control at the expense of generalizability. All research is not equal in terms of methodological rigor, generalizability, level of control, and replicability. Indicators of strong research include clear and specific research questions, a rigorous research design, clear interpretations of results, transparent reporting, and clear methodology that allows for successful replication (Grand *et al.*, 2018).

Practice

Practice informs us through direct experience and observations as well as the experience of others (e.g., case studies, testimonials; Audi, 2011). It may also come in the form of practitioner reports, such as benchmarking surveys assessing the popularity of various assessment and development

practices. These accounts can be quite engaging and persuasive when they come from highly experienced and respected sources, but they can be context-specific and limited in generalizability.

Commonly, observations of practice in organizations lead to consensus around “best practices” that other organizations should implement. A “best practice” is a practice that has functioned successfully in one organization and has the potential to elicit success in another (Serrat, 2017). Its “best” status may be defined by a constellation of factors, including whether the practice is backed by organizational data and how those data were collected. Best practices are often supported by empirical research and aligned with relevant theory. However, assessment and development practices are not required to undergo peer-review processes (like published research) to be implemented; thus, one must evaluate biases that may exist (e.g., conflicts of interest; Lowman & Cooper, 2018). Altogether, knowledge can be acquired through understanding, identifying, and disseminating organizational best practices.

Induction

Practice can uniquely serve as a way of knowing through induction, which can occur during observations of practice. Induction begins with observations of phenomena which can amass into general premises and is in contrast to deduction, in which general premises are used to formulate specific hypotheses. General premises that inductively arise from observations of practice can be formulated into formal theory (Locke, 2007), but such formalization is not a requirement for induction to provide contextualized information relevant for working in the field.

In this way, induction occurring during observations of practice can generate applicable knowledge for those working in the organization. For example, trait activation theory suggests that assessors’ ratings on different dimensions are more likely to converge when the observed behaviors relate to the same underlying trait (Lievens et al., 2006); however, an organization might find similar levels of convergence in ratings on dimensions which have similar and different underlying traits. In this instance, detailed observation of the assessment center in practice can help practitioners understand factors impacting the convergence of ratings for dimensions influenced by dissimilar underlying traits. Then, such factors can be incorporated into future research designs, and subsequently, theory.

Traditionally, organizational scholarship has disparaged making conclusions about organizational practices based on case studies, instead favoring deductive methods, where hypotheses are formed and tested empirically based on prior theory and empirical research (Platt, 1964; Popper, 2003). However, the value of more inductive approaches, where knowledge is acquired through such observation has been shown to be appropriate as well in some contexts, and at times even groundbreaking (Locke, 2007). Indeed, it was through such inductive approaches that key advances in psychological knowledge have come about (e.g., social-cognitive theory; Bandura, 1986; goal setting theory; Locke & Latham, 1990). Woo et al. (2017) advocate for inductive knowledge acquisition when the knowledge seeker begins with a clear purpose, exploits available data, remains flexible and thinks outside of the box, engages in collaborative information-sharing, seeks to replicate and cross-validate conclusions, and reports the observation collection process transparently.¹ As such, observations of practice are essential to the acquisition of knowledge, and can be used systematically to inform assessment and development theory, empirical research, and subsequent practice.

Advantages and disadvantages of practice

Practice contributes to understanding in a unique way as compared to theory and empirical research. Individuals and organizations interpret information differently based on their perception of the information, oftentimes informed by their social context (Berger & Luckmann, 1966).

¹This form of knowledge building is not unlike the data-driven, big-data methods described above, of which Woo et al.’s (2017) best practices would also apply.

Consequently, theories and research may suggest that a phenomenon will result in a certain outcome in an organization, but in reality, the predicted outcome does not occur. In this way, observations of practices in organizations can contribute a way of knowing. At the same time, unsuccessful organizational practices can also provide fodder for new scientific advances. Strengthening partnerships among organizations and scientists can lead to research and theoretical advances that inform best practices and ultimately increase the robustness of the science (Grand *et al.*, 2018). On the other hand, knowledge that arises from observations of practice may not generalize to dissimilar organizational contexts and may lack explanatory power. In sum, observations of practice can provide contextualized knowledge, but is important to understand that this knowledge may not generalize to all contexts and the mechanisms which explain the phenomenon may be unclear.

The value of monitoring all segments

Broadly, the goal of practitioners is to make informed decisions based on observations and accessible research, and the goal of scholars is to use the extant literature alongside practitioners' observations to strengthen what we know; in turn, allowing practitioners to make better, more informed decisions. For scholars, this means seeking out critical evidence of what works in practice and conducting research across laboratory and field settings to address gaps, explain inconsistent findings, and develop sound theory applicable to real-world organizations. The premise of this article is that a field can have the most certainty about a piece of knowledge if evidence from theory, empirical research, and practice converge – an implicit truth that, by making explicit, we believe can provide a clearer path forward. Lacking this level of support, converging evidence from two segments may provide partial guidance. Importantly, no single segment alone can provide complete certainty.

Table 1 provides a more detailed account of sources of assessment center and training knowledge from theory, empirical research, and practice. Specifically, it provides a summary of “how we know what we know” in each content area, the primary sources where the knowledge originated, and the key secondary sources that integrate and propagate knowledge from the primary sources. We recommend that those seeking to learn more about a specific content area should evaluate knowledge offered from theory, empirical research, and practice using a format similar to Table 1, to inform the research or practice they conduct.

Space limitations limit us from walking through a complete epistemology of these areas. However, in the following sections, we demonstrate how using our epistemology provides a framework for processing what is known in the areas of assessment centers and training—which together, account for a wide swath of assessment and development practices. In doing so, we do not intend to present new information on these topics, but rather, we aim to illustrate how our epistemology provides a framework for organizing exactly what is and is not known in a given area, as well as perspective that highlights how to best move forward in advancing new knowledge.

Theory and empirical research converge (but not Practice)

When theory and research converge with little evidence that practice has followed, it could mean undocumented practice is underway, or that sound theory and research are not well understood or appreciated in practice and scholars need to improve translational communication (Banks & Murphy, 1985). The alignment of strong theory and research provides a foundation for continued research, but it may not prove as useful when it is not utilized in practice.

For example, theory and empirical research on assessment centers show that assessors can only reliably differentiate three to five performance dimensions (Gaugler & Thornton, 1989). This aligns with information processing theory's assertion that individuals can only hold a limited amount of information in their working memory without making errors (Lachman & Butterfield, 1979). Therefore, assessors who rate fewer performance dimensions are more accurate in their

Table 1. Sources of Knowledge from Theory, Empirical Research, and Practice on Assessment Centers and Training

Intervention	Theory		Research		Practice	
	Comprehensive Secondary Sources	Primary Sources	Comprehensive Secondary Sources	Primary Sources	Comprehensive Secondary Sources	Primary Sources
Assessment Center Method	<p>Funder (2012) Used the Realistic Accuracy Model (RAM) to explain the factors that affect the accuracy of personality judgments. The RAM specifies that relevant information about personality must be observable, detected by a judge, and used properly to make an accurate personality judgment.</p> <p>Kleinmann & Ingold (2019) Provided a conceptual overview of (a) the assessor rating process, (b) factors that influence assessor performance, and (c) the relationship between the exercise, assessee, and assessor.</p> <p>Thornton & Lievens (2019) Explained the theoretical principles that guide the development of assessment centers and how to use theory to guide decisions made when developing an assessment center.</p>	<p>Lievens et al. (2006) Applied trait activation theory to the assessment center context and found greater convergence in ratings of behaviors related to the same trait versus different traits.</p> <p>Schollaert, & Lievens (2011, 2012) Applying trait activation theory, they proposed that implementing certain situational stimuli in an assessment center may lead to greater behavioral observability. Results showed that role player-prompts, but not specific exercise instructions, increased observability of behavior.</p>	<p>Thornton et al. (2015) Summarized assessment center research and detailed how knowledge gained from this research can inform how organizations create assessment centers to align with their talent management strategies.</p> <p>Kleinmann & Ingold (2019) Offered a review of the empirical findings on (a) the assessor rating process, (b) factors that influence assessor performance and (c) the relationship between the exercise, assessee, and assessor.</p>	<p>Bray & Grant (1966) Described the process and results of the Bell System’s Management Progress study which used an assessment center to assess the dimensions of young managers and predict their future performance.</p> <p>Schleicher et al. (2002) Results established that frame-of-reference assessor training significantly improved the reliability and validity of assessor scores.</p>	<p>Povah (2011) Provided an overview of survey results on international assessment centers. Summarized information such as the dimensions assessed, types of exercises, assessor training content, and feedback processes.</p> <p>International Task Force on Assessment Center Guidelines [ITFACG] (2015) Detailed professional guidelines that organizations should abide by when implementing assessment centers and emphasized ethical considerations.</p> <p>Schlebusch & Roodt (2019) Provided practical recommendations on the development of assessment centers as well as an overview of the history of assessment centers in South Africa.</p>	<p>Birri & Melcher (2011) Case study detailed Credit Suisse’s use of assessment centers for Human Capital Management purposes.</p> <p>Rupp et al. (2006) Provided instructions for validating developmental assessment centers and evidence for the validity of developmental assessment centers.</p>

(Continued)

Table 1. (Continued)

Intervention	Theory		Research		Practice	
	Comprehensive Secondary Sources	Primary Sources	Comprehensive Secondary Sources	Primary Sources	Comprehensive Secondary Sources	Primary Sources
Training	<p>Burke & Hutchins (2007) Provided an integrative review of the factors impacting transfer of training across areas of psychology, management, human resources, training, and learning. Evaluated theory on transfer by evaluating research progress on learner characteristics, intervention design and delivery, and contextual influences.</p> <p>Noe et al. (2014) Offered themes from the learning literature demonstrating how learning theory impacts organizational objectives in the modern workplace, particularly around developing human capital resources.</p> <p>Tannenbaum et al. (1993) Presented an integrated framework depicting the complex process of training effectiveness and all of the variables that might influence</p>	<p>Bell & Kozlowski (2008) Examined how different training designs impact learning and transfer via cognitive, emotional, and motivational processes. Used self-regulation theory to propose and test an integrated model of active learning.</p> <p>Dweck (1986) Used theory in developmental psychology to describes the motivational processes that influence how people acquire and use knowledge and skills. Presented a model suggesting that goal orientations shape reactions to success and failure, defining adaptive and maladaptive motivational patterns to learn.</p> <p>Kanfer & Ackerman (1989) Presented and tested a theory describing how ability and</p>	<p>Arthur et al. (2003) Meta-analysis which identified the design and evaluation features associated with training effectiveness.</p> <p>Bell et al. (2017) Reviewed 100 years of training and development research to highlight what we know and where future research should focus.</p> <p>Bisbey et al. (2021) Provided a summary of research regarding the specifics behind effective design, delivery, implementation, and evaluation of workplace training.</p>	<p>Keith & Frese (2008) Uncovered evidence that encouraging learners to make mistakes in training is a more effective way to promote transfer than error-avoidant methods.</p> <p>Mathieu et al. (1992) Tested a model of individual and situational influences on training effectiveness.</p> <p>Rouiller & Goldstein (1993) Uncovered the important role of organizational climate on transfer behaviors.</p>	<p>Cascio (2019) Examined current trends in the 21st century workplace and presented a look ahead to propose directions for research and practice.</p> <p>Ford (2021) Provided both depth and breadth in reviewing theory, research, and applied experience to inform learning and development in practice. Included practice exercises.</p> <p>Salas et al. (2012) Provided succinct instructions and evidence-based principles for designing, delivering, and implementing team training interventions</p>	<p>ATD Annual State-of-the-Industry Reports (2020) An annual report published by the Association for Talent Development that described the results of a broad survey on talent development trends, spending, and activities.</p> <p>PayScale (2016) Industry report which detailed the skills gaps of recent college graduates, highlighting the skills that are necessary and lacking in the 21st century workplace.</p>

(Continued)

Table 1. (Continued)

Intervention	Theory		Research		Practice	
	Comprehensive Secondary Sources	Primary Sources	Comprehensive Secondary Sources	Primary Sources	Comprehensive Secondary Sources	Primary Sources
	different levels of this process.	motivation interact during skill acquisition. Suggested that ability and task demands impact attentional resources, which direct self-regulation and motivation required to learn. Kraiger et al. (1993) Presented a theoretically based model of training evaluation considering three different types of learning: affective, skill-based, and cognitive.				

ratings (Thornton *et al.*, 2015). Despite this evidence, operational ACs often assess far more dimensions (i.e., 10–12, and even up to 20; Eurich *et al.*, 2009), demonstrating how organizations have yet to adapt practice supported by theory and research.

A similar example can be found in the training literature. Bell and Kozlowski (2008) advanced a theoretical model and empirical validation explaining how individual differences and training design interact to affect learning and the transfer of knowledge back to the job. Their model has been substantiated and built upon by many researchers (e.g., Blume *et al.*, 2019), demonstrating the importance of taking steps before, during, and after training to support transfer. Nonetheless, a lack of transfer continues to be a key issue in practice regardless of how effective training programs are at facilitating learning; and organizations commonly choose not to incorporate design characteristics to maximize transfer (Velada *et al.*, 2007). There are many reasons for this, including a lack of time, accountability, evaluation efforts, and knowledge of best practices (Hutchins *et al.*, 2010; Longnecker, 2004). Further, research insights are not always applicable in practice (Baldwin *et al.*, 2017), emphasizing a need for translational work with actionable research findings to improve (e.g., training) outcomes, including tools and guidelines that consider practical constraints (e.g., to supporting training transfer; Hughes *et al.*, 2018).

In essence, these examples illustrate the well-known “science-practice gap” (Rynes *et al.*, 2001; Tkachenko *et al.*, 2017). Identifying these types of divides can signal the need to forge partnerships among academics and practitioners to develop knowledge, and to disseminate that knowledge in accessible and actionable formats. A classic example of this lies in the history of the assessment center method. Assessment centers, having been widely successful in the selection of intelligence and military personnel during World War II (MacKinnon, 1977), led AT&T to carry out a longitudinal study of the ability for assessment center ratings to predict the career progression of managers. The work involved a partnership between researchers and consultants, including Douglas Bray and William Byham, respectively. After noting the strong predictive validity of the method, results and techniques were shared openly, not only with the scholarly community through the publication of Bray and Grant’s (1966) article in *Psychological Monographs*, but also Byham’s (1970) practice-focused article published in *Harvard Business Review*. Indeed, it was this sort of transparent science-practice collaboration that led to a huge surge in demand for assessment center consultation, culminating in the founding of the firm Development Dimensions International (Thornton & Rupp, 2006).

Theory and practice converge (but not Research)

Next are instances where theory and practice converge but empirical research has either been limited or unreported. One possible explanation is that practitioners are satisfied with a practice and see no need for supportive research (or have no reason to be concerned about unsupportive research). Also, practitioners may not have had opportunities to engage in meaningful research. Another scenario would be that local, proprietary research has been conducted, but not presented publicly. Indeed, research can lag when organizations pioneer into new areas (e.g., big data; Tonidandel *et al.*, 2016). In these cases, theory may explain successful practice while (public facing) empirical validation awaits.

There exist multiple assessment center practices supported by theory but not empirically tested. For example, theory and practice suggest that motivation, cognitive understanding, and experience are related to assessee performance (e.g., Guidry *et al.*, 2013). However, empirical research has yet to explore these issues in depth. Similarly, the use of virtual ACs has greatly increased within practice, with a number of conceptual and theory-based papers written about their use (e.g., Lanik, 2011; Reynolds & Rupp, 2010; Rupp *et al.*, 2008), despite limited reliability and validity evidence to support this modality (for related examples, see Arthur *et al.*, 2014; Illingworth *et al.*, 2015; Morelli *et al.*, 2014). Research is needed that investigates these

theory-backed practices to establish their validity and generalizability across uses (selection, development), organizational levels (entry, management), and industries.

Similarly, the topic of training sustainability, including the need for refresher training (Lazarra et al., 2021), illustrates this segment of the epistemology. Both theory on skill decay over time and observations of decay in practice suggest that refresher training is often necessary to maintain an appropriate level of expertise. Meta-analyses suggest that skills decay with nonuse (Arthur et al., 1998). However, research lends limited insight demonstrating the specifics of what works and when refresher training might be necessary (Lazarra et al., 2021). Needs vary across organizations and skill types, as well as knowledge domains, new developments, and amount of practice in the performance context. Some skills may need to be refreshed less often because they are practiced regularly, while others that are important but used infrequently (e.g., emergency procedures) likely need refreshing sooner. Research is needed that integrates theories such as those on skill decay (Arthur et al., 1998) to create and test frameworks for training sustainability.

Research and practice converge (but not Theory)

When empirical research and practice converge without theoretical explanation, the argument could be made that theory does not matter or explanations are not needed. However, this “black box” or “dustbowl” empiricism can be problematic when issues arise in practice and there is no clear understanding of the mechanisms driving the effectiveness of various assessment and development techniques (Pam, 2020). For instance, multiple innovations within assessment and development have been put into practice while lacking explanatory theory (Lievens & Thornton, 2005); including speed assessments (Herde & Lievens, 2020), automatic scoring of job candidate essays and interviews (Campion et al., 2016; Chen et al., 2022), and asynchronous assessment (Lukacik et al., 2022). These modern practices may seem more resource-efficient than traditional methods. However, we lack theory necessary for understanding the mechanisms underlying their efficacy.

Similarly, the training field lacks an overall, cohesive theory that details the role of training within larger talent development and organizational effectiveness frameworks. For instance, considering the various ways employees develop expertise, it is currently unclear how formal and informal learning efforts might interact and how different channels might be leveraged to maximize benefits. Informal, on-the-job learning has become a key learning pathway in practice (making up 70-90% of all learning activities) and is effective at improving learning and performance (ATD, 2020; Cerasoli et al., 2018). Traditionally, organizations have not placed much attention on informal learning, but employees are now seeking these opportunities and may benefit from structure (ATD, 2020; Cerasoli et al., 2018). Ideally, this structure would be grounded in theory and backed by research on what employees can gain from informal learning in conjunction with training and other talent management interventions.

Conclusion

In this paper, we introduced a unique three-pronged epistemology for determining what is known about assessment and development. We then showed several topics within the assessment center and training areas where two sources of knowledge converge, but one source is lacking. Shockingly, we were not able to locate an example of complete epistemological convergence (i.e., the grey ABC segment of Figure 1). This should serve as a wake-up call to all those working in the assessment and development space, that we must do a better job working together to collect credible insights from theory, research, and practice, followed by the careful and systematic assessment of convergence to come to conclusions on what we confidently “know” about any given area.

This epistemological approach could be applied to other complex workplace interventions, as well. For example, pay and benefit plans vary considerably, and differing sources (i.e., theory, research, and practice) recommend different options (e.g., hourly vs. salary pay, incentives, profit

sharing, early retirement, cafeteria-style benefit plans; Martocchio, 2020). Likewise, this epistemological approach could advance knowledge involving interventions to meet current and future workplace challenges, such as hybrid work arrangements; creating organizational cultures that value diversity and inclusion; and programs that support employee mental health and well-being.

Declaration of Interest. None.

References

- American Psychological Association.** APA Dictionary of Psychology (2nd edn.) 2015. Retrieved July 19, 2022, from <https://dictionary.apa.org/research>.
- American Psychological Association.** APA Style Journal Article Reporting Standards. 2020. Retrieved July 19, 2022, from <https://apastyle.apa.org/jars>.
- Arthur, W., Jr., Bennett, W., Jr., Edens, P. S., & Bell, S. T.** (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. *Journal of Applied Psychology*, *88*(2), 234–245. Doi: [10.1037/0021-9010.88.2.234](https://doi.org/10.1037/0021-9010.88.2.234).
- Arthur, W., Jr., Bennett, W., Jr., Stanush, P. L., & McNelly, T. L.** (1998). Factors that influence skill decay and retention: A quantitative review and analysis. *Human Performance*, *11*(1), 57–101. Doi: [10.1207/s15327043hup1101_3](https://doi.org/10.1207/s15327043hup1101_3).
- Arthur, W., Jr., Doverspike, D., Munoz, G. J., Taylor, J. E., & Carr, A. E.** (2014). The use of mobile devices in high-stakes remotely delivered assessments and testing. *International Journal of Selection and Assessment*, *22*, 113–123.
- Association for Talent Development [ATD]** (2020). *State of the industry: Talent development benchmarks and trends*. ATD Research.
- Audi, R.** (2011). *Epistemology: A contemporary introduction to the theory of knowledge* (3rd edn.) Routledge.
- Bacharach, S. B.** (1989). Organizational theories: Some criteria for evaluation. *The Academy of Management Review*, *14*(4), 496. Doi: [10.2307/258555](https://doi.org/10.2307/258555).
- Baldwin, T. T., Ford, J. K., & Blume, B. D.** (2017). The state of transfer of training research: Moving toward more consumer-centric inquiry. *Human Resource Development Quarterly*, *28*(1), 17–28. Doi: [10.1002/hrdq.21278](https://doi.org/10.1002/hrdq.21278).
- Bandura, A.** (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.
- Banks, C. G., & Murphy, K. R.** (1985). Toward narrowing the research-practice gap in performance appraisal. *Personnel Psychology*, *38*(2), 335–345. Doi: [10.1111/j.1744-6570.1985.tb00551.x](https://doi.org/10.1111/j.1744-6570.1985.tb00551.x).
- Barends, E., & Rousseau, D. M.** (2018). *Evidence-based management: How to use evidence to make better organizational decisions*. Kogan Page.
- Barends, E., Rousseau, D. M., & Briner, R. B.** (2014). *Evidence-based management: The basic principles*. The Center for Evidence-Based Management.
- Bell, B. S., & Kozlowski, S. W. J.** (2008). Active learning: Effects of core training design elements on self-regulatory processes, learning, and adaptability. *Journal of Applied Psychology*, *93*(2), 296–316. Doi: [10.1037/0021-9010.93.2.296](https://doi.org/10.1037/0021-9010.93.2.296).
- Bell, B. S., Tannenbaum, S. I., Ford, J. K., Noe, R. A., & Kraiger, K.** (2017). 100 years of training and development research: What we know and where we should go. *Journal of Applied Psychology*, *102*(3), 305–323. Doi: [10.1037/apl0000142](https://doi.org/10.1037/apl0000142).
- Benjamin, L. T., Jr., & Baker, D. B.** (2000). Boulder at 50: Introduction to the section. *American Psychologist*, *55*(2), 233–254.
- Berger, P. L., & Luckmann, T.** (1966). *The social construction of reality: A treatise in the sociology of knowledge*. Anchor Books.
- Berry, W. D., & Sanders, M. S.** (2018). *Understanding multivariate research: A primer for beginning social scientists*. Routledge.
- Birri, R., & Melcher, A.** (2011). Building a talent for talent. In N. Povah, & G. C. Thornton III (Eds.), *Assessment centres and global talent management* (pp. 175–192). Gower.
- Bisbey, T. M., Grossman, R., Pantan, K., Coultas, C., & Salas, E.** (2021). Design, delivery, evaluation, and transfer of effective training systems. In G. Salvendy, & W. Karwowski (Eds.), *Handbook of human factors and ergonomics* (pp. 414–433). John Wiley & Sons.
- Blume, B. D., Ford, J. K., Surface, E. A., & Olenick, J.** (2019). A dynamic model of training transfer. *Human Resource Management Review*, *29*(2), 270–283. Doi: [10.1016/j.hrmr.2017.11.004](https://doi.org/10.1016/j.hrmr.2017.11.004).
- Bray, D. W., & Grant, D. L.** (1966). The assessment center in the measurement of potential for business management. *Psychological Monographs: General and Applied*, *80*(17), 1–27. Doi: [10.1037/h0093895](https://doi.org/10.1037/h0093895).
- Burke, L. A., & Hutchins, H. M.** (2007). Training transfer: An integrative literature review. *Human Resource Development Review*, *6*(3), 263–296. Doi: [10.1177/1534484307303035](https://doi.org/10.1177/1534484307303035).
- Byham, W. C.** (1970). Assessment centers for spotting future managers. *Harvard Business Review*, *48*, 150–164.
- Campion, M. C., Campion, M. A., Campion, E. D., & Reider, M. H.** (2016). Initial investigation into computer scoring of candidate essays for personnel selection. *Journal of Applied Psychology*, *101*(7), 958–975. Doi: [10.1037/apl0000108](https://doi.org/10.1037/apl0000108).

- Cascio, W. F. (2019). Training trends: Macro, micro, and policy issues. *Human Resource Management Review*, *29*(2), 284–297. Doi: [10.1016/j.hrmr.2017.11.001](https://doi.org/10.1016/j.hrmr.2017.11.001).
- Cerasoli, C. P., Alliger, G. M., Donsbach, J. S., Mathieu, J. E., Tannenbaum, S. I., & Orvis, K. A. (2018). Antecedents and outcomes of informal learning behaviors: A meta-analysis. *Journal of Business and Psychology*, *33*, 203–230. Doi: [10.1007/s10869-017-9492-y](https://doi.org/10.1007/s10869-017-9492-y).
- Chen, K., Niu, M., & Chen, Q. (2022). A hierarchical reasoning graph neural network for the automatic scoring of answer transcriptions in video job interviews. *International Journal of Machine Learning and Cybernetics*, *13*, 2507–2517. Doi: [10.1007/s13042-022-01540-8](https://doi.org/10.1007/s13042-022-01540-8).
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*, vol. 351. Houghton Mifflin.
- Cucina, J. M., Hayes, T. L., Walmsley, P. T., & Martin, N. R. (2014). It is time to get medieval on the overproduction of pseudotheory: How Bacon, 1267 and Alhazen, 1021 can save Industrial-Organizational Psychology. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *7*(3), 356–364. Doi: [10.1111/iops.12163](https://doi.org/10.1111/iops.12163).
- Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*, Retrieved from, <https://www.reuters.com/article/us-amazon-comjobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*(10), 1040–1048. Doi: [10.1037/0003-066X.41.10.1040](https://doi.org/10.1037/0003-066X.41.10.1040).
- Eurich, T. L., Krause, D. E., Cigularov, K., & Thornton, G. C. (2009). Assessment centers: Current practices in the United States. *Journal of Business and Psychology*, *24*, 387. Doi: [10.1007/s10869-009-9123-3](https://doi.org/10.1007/s10869-009-9123-3).
- Folger, R., & Stein, C. (2017). Abduction 101: Reasoning processes to aid discovery. *Human Resource Management Review*, *27*(2), 306–315. Doi: [10.1016/j.hrmr.2016.08.007](https://doi.org/10.1016/j.hrmr.2016.08.007).
- Ford, J. K. (2021). *Learning in organizations: An evidence-based approach*. Taylor & Francis.
- Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, *21*(3), 177–182. Doi: [10.1177/0963721412445309](https://doi.org/10.1177/0963721412445309).
- Gaugler, B. B., & Thornton, G. C., III (1989). Number of assessment center dimensions as a determinant of assessor accuracy. *Journal of Applied Psychology*, *74*(4), 611–618. Doi: [10.1037/0021-9010.74.4.611](https://doi.org/10.1037/0021-9010.74.4.611).
- Grand, J. A., Rogelberg, S. G., Allen, T. D., Landis, R. S., Reynolds, D. H., Scott, J. C., Tonidandel, S., & Truxillo, D. M. (2018). A systems-based approach to fostering robust science in Industrial-Organizational Psychology. *Industrial and Organizational Psychology*, *11*(1), 4–42. Doi: [10.1017/iop.2017.55](https://doi.org/10.1017/iop.2017.55).
- Groves, R. M. (2011). Three eras of survey research. *Public Opinion Quarterly*, *75*(5), 861–871. Doi: [10.1093/poq/nfr057](https://doi.org/10.1093/poq/nfr057).
- Guidry, B., Rupp, D., & Lanik, M. (2013). Tracing cognition with assessment center simulations: Using technology to see in the dark. In M. Fetzer, & K. Tuzinsk (Eds.), *Simulations for Personnel Selection* (pp. 231–258). Springer.
- Hambrick, D. C. (2007). The field of management's devotion to theory: Too much of a good thing? *Academy of Management Journal*, *50*(6), 1346–1352. Doi: [10.5465/amj.2007.28166119](https://doi.org/10.5465/amj.2007.28166119).
- Herde, C. N., & Lievens, F. (2020). Multiple speed assessments: Theory, practice, and research evidence. *European Journal of Psychological Assessment*, *36*(2), 237–249. Doi: [10.1027/1015-5759/a000512](https://doi.org/10.1027/1015-5759/a000512).
- Hickman, L., Tay, L., & Woo, S. E. (2019). Validity evidence for off-the-shelf language-based personality assessment using video interviews: Convergent and discriminant relationships with self and observer ratings. *Personnel Assessment and Decisions*, *5*(3), 12–20. Doi: [10.25035/pad.2019.03.003](https://doi.org/10.25035/pad.2019.03.003).
- Huffcutt, A. I. (2004). Research perspectives on meta-analysis. In S. G. Rogelberg (Eds.), *Handbook of research methods in industrial and organizational psychology* (pp. 198–215). Blackwell Publishing.
- Hughes, A. M., Zajac, S., Spencer, J. M., & Salas, E. (2018). A checklist for facilitating training transfer in organizations. *International Journal of Training and Development*, *22*(4), 334–345. Doi: [10.1111/ijtd.12141](https://doi.org/10.1111/ijtd.12141).
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage.
- Hutchins, H. M., Burke, L. A., & Berthelsen, A. M. (2010). A missing link in the transfer problem? Examining how trainers learn about training transfer. *Human Resource Management*, *49*(4), 599–618. Doi: [10.1002/hrm.20371](https://doi.org/10.1002/hrm.20371).
- Illingworth, A. J., Morelli, N. A., Scott, J. C., & Boyd, S. L. (2015). Internet-based, unproctored assessments on mobile and non-mobile devices: Usage, measurement equivalence, and outcomes. *Journal of Business Psychology*, *30*, 325–343. Doi: [10.1007/s10869-014-9363-8](https://doi.org/10.1007/s10869-014-9363-8).
- International Task Force on Assessment Center Guidelines [ITFACG] (2015). Guidelines and ethical considerations for assessment center operations. *Journal of Management*, *41*(4), 1244–1273. Doi: [10.1177/0149206314567780](https://doi.org/10.1177/0149206314567780).
- Jackson, S. E., & Schuler, R. S. (1995). Understanding human resource management in the context of organizations and their environments. *Annual Review of Psychology*, *46*(1), 237–264. Doi: [10.1146/annurev.ps.46.020195.001321](https://doi.org/10.1146/annurev.ps.46.020195.001321).
- Johns, G. (1993). Constraints on the adoption of psychology-based personnel practices: Lessons from organizational innovation. *Personnel Psychology*, *46*(3), 569–592. Doi: [10.1111/j.1744-6570.1993.tb00885.x](https://doi.org/10.1111/j.1744-6570.1993.tb00885.x).
- Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/ aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, *74*(4), 657–690. Doi: [10.1037/0021-9010.74.4.657](https://doi.org/10.1037/0021-9010.74.4.657).

- Keith, N., & Frese, M. (2008). Effectiveness of error management training: A meta-analysis. *Journal of Applied Psychology*, *93*(1), 59–69. Doi: [10.1037/0021-9010.93.1.59](https://doi.org/10.1037/0021-9010.93.1.59).
- Klein, K. J., & Kozlowski, S. W. J. Eds (2000). *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*. Jossey-Bass.
- Kleinmann, M., & Ingold, P. V. (2019). Toward a better understanding of assessment centers: A conceptual review. *Annual Review of Organizational Psychology and Organizational Behavior*, *6*, 349–372. Doi: [10.1146/annurev-orgpsych-012218-014955](https://doi.org/10.1146/annurev-orgpsych-012218-014955).
- Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training. *Journal of Applied Psychology*, *78*(2), 311–328. Doi: [10.1037/0021-9010.78.2.311](https://doi.org/10.1037/0021-9010.78.2.311).
- Lachman, R., Lachman, J. L., & Butterfield, E. C. (1979). *Cognitive psychology and information processing: An introduction* (1st edn.) Psychology Press.
- Lanik, M. (2011). *Breaking the tradition: AC 2.0. Presentation at the 31st*. Assessment Center Study Group Conference, Somerset West, South Africa.
- Lazzara, E. H., Benishek, L. E., Hughes, A. M., Zajac, S., Spencer, J. M., Heyne, K. B., Rogers, J. E., & Salas, E. (2021). Enhancing the organization's workforce: Guidance for effective training sustainment. *Consulting Psychology Journal: Practice and Research*, *73*(1), 1–26. Doi: [10.1037/cpb0000185](https://doi.org/10.1037/cpb0000185).
- LeBreton, J. M., Scherer, K. T., & James, L. R. (2014). Corrections for criterion reliability in validity generalization: A false prophet in a land of suspended judgment: Corrections for criterion. *Industrial and Organizational Psychology*, *7*(4), 478–500. Doi: [10.1111/iops.12184](https://doi.org/10.1111/iops.12184).
- Lievens, F., Chasteen, C. S., Day, E. A., & Christiansen, N. D. (2006). Large-scale investigation of the role of trait activation theory for understanding assessment center convergent and discriminant validity. *Journal of Applied Psychology*, *91*(2), 247–258. Doi: [10.1037/0021-9010.91.2.247](https://doi.org/10.1037/0021-9010.91.2.247).
- Lievens, F., & Thornton, G. C. (2005). Assessment Centers: Recent Developments in Practice and Research. In A. Evers, N. Anderson, & O. Voskuilj (Eds.), *The Blackwell handbook of personnel selection* (pp. 243–264).
- Locke, E. A. (2007). The case for inductive theory building. *Journal of Management*, *33*(6), 867–890. Doi: [10.1177/0149206307307636](https://doi.org/10.1177/0149206307307636).
- Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Prentice Hall.
- Longnecker, C. O. (2004). Maximizing transfer of learning from management education programs: Best practices for retention and application. *Development and Learning in Organizations*, *18*(4), 4–6. Doi: [10.1108/1477280410544538](https://doi.org/10.1108/1477280410544538).
- Lowman, R. L., & Cooper, S. E. (2018). *The ethical practice of consulting psychology*. American Psychological Association.
- Lukacik, E. R., Bourdage, J. S., & Roulin, N. (2022). Into the void: A conceptual model and research agenda for the design and use of asynchronous video interviews. *Human Resource Management Review*, *32*(1), 100789. Doi: [10.1016/j.hrmr.2020.100789](https://doi.org/10.1016/j.hrmr.2020.100789).
- MacKinnon, D. W. (1977). From selecting spies to selecting managers—the OSS assessment program. In J. J. Moses, & W. C. Byham (Eds.), *Applying the assessment center method* (pp. 13–30). Pergamon Press.
- Martocchio, J. J. (2020). *Strategic compensation: A human resource management perspective* (10th edn.) Pearson.
- Mathieu, J. E., Tannenbaum, S. I., & Salas, E. (1992). Influences of individual and situational characteristics on measures of training effectiveness. *Academy of Management Journal*, *35*(4), 828–847. Doi: [10.5465/256317](https://doi.org/10.5465/256317).
- McAbee, S. T., Landis, R. S., & Burke, M. I. (2017). Inductive reasoning: The importance of big data. *Human Resource Management Review*, *27*(2), 277–290. Doi: [10.1016/j.hrmr.2016.08.005](https://doi.org/10.1016/j.hrmr.2016.08.005).
- McLelland, C. V. (2006). *The nature of science and the scientific method*. The Geological Society of America.
- Morelli, N. A., Mahan, A., & Illingworth, J. (2014). Establishing the measurement equivalence of online selection assessments delivered on mobile versus nonmobile devices. *International Journal of Selection and Assessment*, *22*(2), 124–138. Doi: [10.1111/ijasa.12063](https://doi.org/10.1111/ijasa.12063).
- Murphy, K. R., Cleveland, J. N., & Hanscom, M. E. (2019). *Performance appraisal and management*. Sage.
- Nealon, J. T., & Giroux, S. S. (2012). *The theory toolbox: Critical concepts for the humanities, arts, and social sciences* (2nd edn.) Rowman & Littlefield Publishers.
- Noe, R. A., Clarke, A. D. M., & Klein, H. J. (2014). Learning in the twenty-first-century workplace. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*, 245–275. Doi: [10.1146/annurev-orgpsych-031413-091321](https://doi.org/10.1146/annurev-orgpsych-031413-091321).
- Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, *78*(4), 679–703. Doi: [10.1037/0021-9010.78.4.679](https://doi.org/10.1037/0021-9010.78.4.679).
- Pam, M. S. (2020). Dustbowl empiricism. *Psychology Dictionary*, Accessed April 3, 2020.
- Parrigon, S., Woo, S. E., Tay, L., & Wang, T. (2017). CAPTION-ing the situation: A lexically-derived taxonomy of psychological situation characteristics. *Journal of Personality and Social Psychology*, *112*(4), 642–681. Doi: [10.1037/pssp0000111](https://doi.org/10.1037/pssp0000111).

- PayScale. 2016 workforce-skills preparedness report. 2016. Retrieved July 26, 2021, from <https://www.payscale.com/data-packages/job-skills>.
- Platt, J. R. (1964). Strong inference. *Science*, *146*(3642), 347–353. Doi: [10.1126/science.146.3642.347](https://doi.org/10.1126/science.146.3642.347).
- Ployhart, R. E., & Bartunek, J. M. (2019). Editors' comments: There is nothing so theoretical as good practice – A call for phenomenal theory. *Academy of Management Review*, *44*, 493–497. Doi: [10.5465/amr.2019.0087](https://doi.org/10.5465/amr.2019.0087).
- Popper, K. R. (2003). *The logic of scientific discovery*. Routledge.
- Povah, N. (2011). A review of recent international surveys into assessment centre practices. In N. Povah, & G. C. Thornton III (Eds.), *Assessment centres and global talent management* (pp. 329–350).
- Reynolds, D. H., & Rupp, D. E. (2010). Advances in technology-facilitated assessment. In J. C. Scott, & D. H. Reynolds (Eds.), *Handbook of workplace assessment: Evidence-based practices for selecting and developing organizational talent* (pp. 609–641). Jossey-Bass.
- Rouiller, J. Z., & Goldstein, I. L. (1993). The relationship between organizational transfer climate and positive transfer of training. *Human Resource Development Quarterly*, *4*(4), 377–390. Doi: [10.1002/hrdq.3920040408](https://doi.org/10.1002/hrdq.3920040408).
- Rousseau, D. M., & Gunia, B. C. (2016). Evidence-based practice: The psychology of EBP implementation. *Annual Review of Psychology*, *67*(1), 667–692. Doi: [10.1146/annurev-psych-122414-033336](https://doi.org/10.1146/annurev-psych-122414-033336).
- Rupp, D. E., & Beal, D. (2007). Checking in with the scientist-practitioner model: How are we doing? *The Industrial-Organizational Psychologist*, *45*(1), 35–40. Doi: [10.1037/e579082011-003](https://doi.org/10.1037/e579082011-003).
- Rupp, D. E., Gibbons, A. M., & Snyder, L. A. (2008). The role of technology in enabling third-generation training and development. *Industrial and Organizational Psychology*, *1*(4), 496–500. Doi: [10.1111/j.1754-9434.2008.00095.x](https://doi.org/10.1111/j.1754-9434.2008.00095.x).
- Rupp, D. E., Gibbons, A. M., Snyder, L. A., Spain, S. M., Woo, S. E., Brummel, B. J., Sims, C. S., & Kim, M. (2006). An initial validation of developmental assessment centers as accurate assessments and effective training interventions. *The Psychologist-Manager Journal*, *9*(2), 171–200. Doi: [10.1207/s15503461tpmj0902_7](https://doi.org/10.1207/s15503461tpmj0902_7).
- Rupp, D. E., Shapiro, D. L., Folger, R., Skarlicki, D. P., & Shao, R. (2017). A critical analysis of the conceptualization and measurement of organizational justice: Is it time for reassessment? *Academy of Management Annals*, *11*(2), 919–959. Doi: [10.5465/annals.2014.0051](https://doi.org/10.5465/annals.2014.0051).
- Rynes, S. L., & Bartunek, J. M. (2017). Evidence-based management: Foundations, development, controversies and future. *Annual Review of Organizational Psychology and Organizational Behavior*, *4*, 235–261. Doi: [10.1146/annurev-orgpsych-032516-113306](https://doi.org/10.1146/annurev-orgpsych-032516-113306).
- Rynes, S. L., Bartunek, J. M., & Daft, R. L. (2001). Across the great divide: Knowledge creation and transfer between practitioners and academics. *Academy of Management Journal*, *44*(2), 340–355. Doi: [10.5465/3069460](https://doi.org/10.5465/3069460).
- Rynes-Weller, S. L. (2012). The research-practice gap in I/O Psychology and related fields: Challenges and potential solutions. In S. W. J. Kozlowski (Eds.), *The Oxford Handbook of Organizational Psychology*. Oxford Academic.
- Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2022). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. *Journal of Applied Psychology*, *107*(11), 2040–2068. Doi: [10.1037/apl0000994](https://doi.org/10.1037/apl0000994).
- Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The science of training and development in organizations: What matters in practice. *Psychological Science in the Public Interest*, *13*(2), 74–101. Doi: [10.1177/1529100612436661](https://doi.org/10.1177/1529100612436661).
- Schlebusch S., & Roodt G. (Eds. 2019). *Assessment centres: Unlocking potential for growth*. 2nd Edition, Knowres.
- Schleicher, D. J., Day, D. V., Mayes, B. T., & Riggio, R. E. (2002). A new frame for frame-of-reference training: Enhancing the construct validity of assessment centers. *Journal of Applied Psychology*, *87*(4), 735–746. Doi: [10.1037/0021-9010.87.4.735](https://doi.org/10.1037/0021-9010.87.4.735).
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*(2), 262–274. Doi: [10.1037/0033-2909.124.2.262](https://doi.org/10.1037/0033-2909.124.2.262).
- Schollaert, E., & Lievens, F. (2011). The use of role player prompts in assessment center exercises. *International Journal of Selection and Assessment*, *19*(2), 190–197. Doi: [10.1111/j.1468-2389.2011.00546.x](https://doi.org/10.1111/j.1468-2389.2011.00546.x).
- Schollaert, E., & Lievens, F. (2012). Building situational stimuli in assessment center exercises: Do specific exercise instructions and role-player prompts increase observability of behavior? *Human Performance*, *25*(3), 255–271. Doi: [10.1080/08959285.2012.683907](https://doi.org/10.1080/08959285.2012.683907).
- Serrat, O. (2017). Identifying and sharing good practices. In O. Serrat (Eds.), *Knowledge solutions: Tools, methods, and approaches to drive organizational performance* (pp. 843–846). Springer.
- Shieber, J. N. (2019). *Theories of knowledge: How to think about what we know*. The Great Courses.
- Snijders, T. A., & Bosker, R. J. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage.
- Suddaby, R. (2014). Editor's comments: Why theory? *Academy of Management Review*, *39*(4), 407–411. Doi: [10.5465/amr.2014.0252](https://doi.org/10.5465/amr.2014.0252).
- Sutton, R. I., & Staw, B. M. (1995). What theory is not. *Administrative Science Quarterly*, *40*(3), 371–384. Doi: [10.2307/2393788](https://doi.org/10.2307/2393788).

- Tannenbaum, S. I., Cannon-Bowers, J. A., & Mathieu, J. E.** (1993). *Factors that influence training effectiveness: A conceptual model and longitudinal analysis (Report 93-011)*. Naval Training Systems Center.
- Thornton, G. C., & Rupp, D. E.** (2006). *Assessment centers in human resource management: Strategies for prediction, diagnosis, and development*. Lawrence Erlbaum Associates.
- Thornton, G. C., III, Rupp, D. E., & Hoffman, B. J.** (2015). *Assessment center perspectives for talent management strategies* (2nd edn.) Routledge/Taylor & Francis Group.
- Thornton, G. C. I. I., & Lievens, F.** (2019). Theoretical principles relevant to assessment center design and implementation. In S. Schlebusch, & G. Roodt (Eds.), *Assessment centres: Unlocking potential for growth* (2nd edn.) Knowres.
- Tkachenko, O., Hahn, H.-J., & Peterson, S. L.** (2017). Research-practice gap in applied fields: An integrative literature review. *Human Resource Development Review*, **16**(3), 235–262. Doi: [10.1177/1534484317707562](https://doi.org/10.1177/1534484317707562).
- Tonidandel, S., King, E., & Cortina, J.** (2016). Big data methods: Leveraging modern data analytic techniques to build organizational science. *Organizational Research Methods*, **21**(3), 525–547. Doi: [10.1177/1094428116677299](https://doi.org/10.1177/1094428116677299).
- Van Iddekinge, C. H., Roth, P. L., Raymark, P. H., & Odle-Dusseau, H. N.** (2012). The criterion-related validity of integrity tests: An updated meta-analysis. *Journal of Applied Psychology*, **97**(3), 499–530. Doi: [10.1037/a0021196](https://doi.org/10.1037/a0021196).
- Velada, R., Caetano, A., Michel, J. W., Lyons, B. D., & Kavanagh, M. J.** (2007). The effects of training design, individual characteristics and work environment on transfer of training. *International Journal of Training and Development*, **11**(4), 282–294. Doi: [10.1111/j.1468-2419.2007.00286.x](https://doi.org/10.1111/j.1468-2419.2007.00286.x).
- Woo, S. E., O’Boyle, E. H., & Spector, P. E.** (2017). Best practices in developing, conducting, and evaluating inductive research. *Human Resource Management, Review*, **27**(2), 255–264. Doi: [10.1016/j.hrmmr.2016.08.004](https://doi.org/10.1016/j.hrmmr.2016.08.004).
- Woo, S. E., Tay, L., & Proctor, R. W.** Eds (2020). *Big data in psychological research*. American Psychological Association.
- Zhou, L., Song, Y., Alterman, V., Liu, Y., & Wang, M.** (2019). Introduction to data collection in multilevel research. In S. E. Humphrey, & J. M. LeBreton (Eds.), *The handbook of multilevel theory, measurement, and analysis* (pp. 225–252). American Psychological Association.

Cite this article: Rupp, D. E., Thornton, G. C., III, Bisbey, T. M., Hoover, A. N., Salas, E., & Murphy, K. R. (2024). An epistemology for assessment and development: How do we know what we know? *Industrial and Organizational Psychology* **17**, 252–268. <https://doi.org/10.1017/iop.2024.3>