Citation for published version:
Ward, P, Schragen, JM, Gore, J, Roth, E, Robert, H & Klein, G 2019, Reflections on the Study of Expertise and Its Implications for Tomorrow's World. in P Ward, JM Schraagen, J Gore & E Roth (eds), *The Oxford Handbook of Expertise: Research & Application.* Oxford University Press.
https://doi.org/10.1093/oxfordhb/9780198795872.013.52

DOI:
10.1093/oxfordhb/9780198795872.013.52

Publication date:
2019

Document Version
Peer reviewed version

Link to publication

Reflections on the Study of Expertise and Its Implications for Tomorrow's World

Paul Ward[1], Jan Maarten Schraagen[2,3], Julie Gore[4], Emilie Roth[5], Robert R. Hoffman[6], & Gary Klein[7]

[1]University of Northern Colorado, USA, [2]TNO, Netherlands, [3]University of Twente, Netherlands, [4]University of Bath, UK, [5]Roth Engineering, USA, [6]Institute for Human and Machine Cognition, USA, & [7]Macrocognition LLC, USA

To appear in Ward, P., Schraagen, J. M., Gore, J., & Roth, E. (Eds.). *Oxford Handbook of expertise: Research & Application.* Oxford, UK: Oxford University Press. Correspondence should be directed to Professor Paul Ward, School of Psychological Sciences, University of Northern Colorado, Greeley, CO 80631. E-mail: paul.ward@unco.edu, Telephone: 970-351-2914

Introduction

There is no doubt that the world continues to change at an increasingly fast pace. For example, what was once considered work, such as the kind of physical activity associated with labor-intensive craft, is often now considered a recreational pastime (e.g., physical *work* outs, lifting weights, etc.). Likewise, what were once considered leisurely pursuits, such as reading, writing, analyzing problems, and trying to understand complex phenomena (which were pursued primarily by society's elite because they could afford to invest time in such pastimes) are now considered work for many of today's citizens. In brief, and generally speaking, where work was once mainly physical, it has become increasingly cognitive. What is more, advances in technology have increased the cognitive complexity of work in many domains. The pace of change in technology continues to accelerate the changing nature of this work and, frequently, the rate at which it is carried out. Hence, what worked in yesteryear may not work today, and is unlikely to work in tomorrow's world—at least for most modern work domains. Accordingly, what it takes to succeed in such work must continue to change. The question is, must our view and definition of expertise, and how we measure it, change too?

In this chapter, we reflect on the themes that emerged from the chapters in this Handbook. In particular, we re-examine the definitions of expertise and the idea that expertise is, in part, about increasing one's cognitive ability to adapt to complexity (see Ward et al., 2018). Whilst we appreciate that this argument has some limitations we argue that it is worthy of further exploration. We then take a look at where we have been, as a community of communities of expertise researchers, and whether we are heading in good directions. Finally, we present some food for thought in terms of future areas of expertise studies that are required to continue to move the field forward and, ultimately, to better prepare individuals to operate effectively in tomorrow's workplace. Playing off of the penultimate chapter by Klein, et al., we begin with a quick look back at the contested War on Expertise and ask whether modern work domains require a new view of expertise.

Why are Experts not always Revered?

Compared to some of the views of expertise presented in this Handbook, some researchers have argued that the value of expertise has changed markedly and that experts are becoming increasingly obsolete. One line of argument suggests that experts have simply failed to keep up with the rate of change in the nature of work, and technology's influence on it. For instance, some have devalued expertise because alternatives that are a product of technological advancements (e.g., machine learning algorithms), can sometimes outperform expert decisions. Others have argued that experts are unnecessary, or even unwelcome in some complex domains (e.g., financial markets; voting polls; geopolitical developments) because they are not always right and are sometimes tragically wrong, which erodes public trust in expertise. The counter argument, however is that experts are essential for successful operations in many (if not all) work domains—as is evidenced when an expert retires—and that many of the successes of *expert substitutes* come about because they are based, at least in part, on expert input. Hence, the apparent fallibilities of expertise may be more a question of how expertise is defined, conceptualized, and measured, and the domain in which expertise is observed, than with experts themselves (Klein et al., this volume).

A second reason why experts are revered less in the current *Zeitgeist*, particularly in non-scientific circles, has to do with a politically-motivated erosion of trust in scientific knowledge. Climate change is the poster child for this polemic. Although expertise is much broader than scientific expertise, the particular example of science and its conceptualization may illustrate different views of expertise as put forth in this Handbook. Basically, science is still viewed by many as the isolated, individualized, pursuit of Truth – the discovery of objective, immutable, undisputable facts. Empirical studies of scientific practices have shown, however, that science is a much more collaborative, networked enterprise in which facts are constructed over long periods of time, requiring a lot of convincing of one's colleagues (Latour & Woolgar, 1979). The subtitle of Latour and Woolgar's book, "The construction of scientific facts," may have opened Pandora's box and has been viewed by some as providing ammunition to critics of the scientific endeavor as a whole: if scientific facts are created, then surely we should not place any more trust in them than in other facts that are likewise created (Kofman, 2018)? The ensuing misunderstandings and political implications have so many repercussions (e.g., on the debate on climate change), that Latour has recently felt obliged to mount a defense of science (Latour, 2018) that relies heavily on the important role of networks in producing and sustaining knowledge. Facts remain robust when they are supported by a common culture, much more so than by their veracity. It is precisely the networked enterprise that makes science valuable.

The various conceptualizations of expertise as discussed in this Handbook mirror the conceptualizations of the scientific endeavor. To some, expertise should be explained from an individualistic stance, as the result of a long process of developing and finetuning of mental processes and representations (Gobet, this volume), with underpinning neural mechanisms (Ullen, de Manzano, & Mosing, this volume). To others, expertise should be explained from a sociological stance (Collins & Evans, this volume), or at least from a situated, enacted, embodied (Baber, this volume), triadic (Flach & Voorhorst, this volume), or tripartite (Pfeiffer, this volume) stance, with the latter two perspectives emphasizing the intimate coupling between person, representation and environment. To be fair, most authors, in their theoretical conceptualizations, propose some form of multifactorial perspective in which expertise is viewed as an adaptation that takes place across various levels simultaneously. Drawing a parallel with the preceding discussion on trust in scientific facts, we may advance the similar notion that trust in expertise remains robust only when expertise is supported by colleagues, peers, collaborators, parents, in short the social environment to which a person has adapted. This is a somewhat underdeveloped conceptualization of expertise that we believe is a promising venue for future research.

In the following, we will focus on expertise as adaptive skill.

Expertise as Adaptive Skill

Structured and predictable environments are more conducive to adaptation, and hence to the development of expertise, than unstructured and unpredictable environments (Shanteau, 1992; Kahneman & Klein, 2009). Second, experts always need to adapt to surprise within their domains of expertise. It is an open question what resources experts may draw upon to be able to continue to adapt to surprise. Most research so far has looked at individual resources, such as types of knowledge representations (Gobet, this volume) or problem-solving strategies (e.g., heuristics, Hoffrage, this volume). Other types of resources, such as team or organizational resources, have been studied to a far lesser extent. Therefore, based on what we know so far, the following discussion on expertise as adaptive skill relies heavily on an individualistic perspective, from a macrocognitive perspective on expertise (see also Hutton, this volume).

In many realms, expertise is defined with respect to routine mastery of a primary task or procedure (for a review of expertise definitions see introductory chapter; Ward et al., this volume), where performance is "outstanding in terms of speed, accuracy, and automaticity of performance" (see Hatano & Inagaki 1984, p. 31). Such definitions of routine expertise may not capture the expert's adaptive capability but they are consistent with current definitions of transfer (see Hoffman, Ward, Feltovich, DiBello, Fiore, & Andrews, 2014). Based on the way in which transfer has been measured traditionally, the research suggests that transfer of expertise is only likely to occur when engaging in similar domain-specific tasks (for a discussion of this issue, see Ward et al., 2018).

Human beings are adaptive, or otherwise they would not survive in changing environments. By definition, then, experts are also adaptive (Hoffman, 1998), but this claim goes beyond the blanket statement of adaptivity as a general human characteristic. It means that experts, by nature of their expertise, have developed specific characteristics that allow them to be more adaptive than non-experts. Experts' adaptiveness however, does not extend to any random domain. As mentioned above, when discussing limits to transfer, it is well-known that expertise is quite domain-specific. Setting aside the fuzzy nature of the boundaries between *domains*, the interesting empirical question is whether experts are more adaptive than non-experts in their domain of expertise. Typically, they have a good conceptual understanding that permits the development and use of a context-sensitive strategy, which allows them to readily identify, both *a priori* and *in situ*, key decision points in a specific course of action and judge when variations of an existing course of action might be in/appropriate. Having a well-indexed conceptual representation of the current situation allows the expert to immediately access opportunities to deviate from the outcome path by selecting, modifying, or generating anew, both situational interpretations and alternative courses of action *on the fly*. Hatano and Inagaki (1984/86) asserted that it is one's conceptual understanding of procedural skill that affords in-event adaptive thinking. We would argue that both conceptual understanding (or sensemaking capabilities) and flexible decision making (or replanning) are two parts of an integrated dynamic system that give rise to successful adaptation in both familiar and new complex contexts within one's domain of expertise.

This characterization of expertise as skilled adaptation to complexity and novelty, at least in their domains of expertise, is consistent with the views of many of the authors of chapters in this Handbook. For instance, Bohle Carbonell and Van Merrienboer (this volume) and Fletcher and Kowal (this volume) defined the adaptive nature of expertise, respectively, as the ability to "perform at a relatively high level in unfamiliar situations" and "successfully solve uncommon, unusually difficult, and/or strategic problems that others cannot." Naikar and Brady (this volume) drew on Rasmussen's (i.e., Rasmussen, 1986; Rasmussen, Pejtersen, & Goodstein, 1994) view of expertise and highlighted the volatility of many human-technological work environments that require adaptation, suggesting that it is about experts being "able to deal successfully with ongoing and significant instability, uncertainty, and unpredictability in their work." This point was echoed by many, including Fletcher and Kowal (this volume) who proclaimed that "unpredictability is a certainty. Dealing with surprise and the unexpected is, then, an inevitable aspect of… expertise." However, they also noted that despite *change* being "difficult to pursue, perilous to conduct, and uncertain in its success" adaptation to change is not always rewarded as a key, or even the main, component of expertise. On this point, Hatano and Inagaki (1984) suggested that in order for adaptive skill to flourish, not only must experts be given the *opportunity to explore* task variations, such exploration of system constraints must be *valued*, and the experts must have the *authority* to explore without reprisal (see also Ward et al., 2018).

These views of the expert as having a flexible and adaptive skill capacity are consistent with a view of expertise presented recently (Ward et al., 2018). Ward et al. (2018) argued that in complex work domains, adaptive skill is the essential ingredient for success—the *conditio sine qua non* of expertise. We extend this argument here by suggesting that the importance of this skill will continue to increase as the societal and human-technological challenges ahead of us proliferate and permeate every aspect of our lives. We therefore need more research to better understand the adaptive character of expertise, and what makes it more apparent and effective in experts rather than novices.

Klein (2011) suggested that in order to become a genuine expert, perhaps an expert of the future, we need to reconceptualize learning. Klein suggests that the *storehouse* metaphor of knowing more and more—that emphasizes putting knowledge *in* to memory—may need to be supplemented by the *snakeskin* metaphor—getting knowledge *out* when needed, and being prepared to shed a particular understanding or course of action for another as the context dictates. In brief, this view promotes a shift away from knowing more toward thinking dynamically, innovatively, and differently—knowing when and when not, and knowing how and why, to generate new solutions on the fly in the face of adversity and anomalies.

One method of developing expertise of this type was presented by Klein (2011) who suggested that we improve performance not just by reducing errors but by increasing insights. Klein laid out several pathways for how we generate insight, which have since been encapsulated in an integrated model of macrocognition (see Figure 1) (see Hoffman & Hancock, 2017, Ward et al., 2018). The integrative D/F + F model combines the Data/Frame (D/F) model of sensemaking and the Flexecution (F) model of adaptive replanning, along with the core concepts in Cognitive Flexibility Theory (Spiro et al., 1992) and Cognitive Transformation Theory (Klein & Baxter, 2006, 2009), all of which are based on data on how experts operate in complex and dynamic environments where adaptation is key to success. In keeping with the integrated model, Ward et al. (2018) proposed a definition of adaptive skill that, we argue here, would have to be central to any future definition of expertise, especially when those experts operate in complex work environments. According to Ward et al., adaptive skill entails:

> Timely changes in understanding, plans, goals, and methods in response to either an altered situation or updated assessment of the ability to meet new demands, that permit successful efforts to achieve intent… or successful efforts to realize alternative statements of intent that are not inconsistent with the initial statement but more likely to achieve beneficial results under changed circumstances (Ward et al., 2018, p. 42).

This definition of adaptive skill captures the expert's requirement to update understanding on the fly in messy, complex, and dynamic environments, which can be thought of as sensemaking (i.e., build the capability to frame and reframe as appropriate, or adaptively reframe). It also captures the iterative and flexible nature of action execution, which can be thought of as a process of flexecution (i.e., build the capability to plan and replan as appropriate, or adaptively replan). Ward et al.'s (2018) definition and the integrated D/F + F model illustrate the interdependence between the sensemaking and flexecution components of adaptation, which can be thought of as managing the tradeoffs within and between components. This tradeoff is necessarily a highly metacognitive and regulatory process of goal evaluation relative to one's intentions and actions in the current context (e.g., see Hoffman, et. al., 2014). Hence, there may be fruitful opportunity to build on the D/F+F integrated model to further explore the adaptive nature of expertise.

As a means to develop adaptive skill, Ward et al. (2018) offered six training principles—based on empirical data from studies of experts in situations where there was a need to adapt—

that we argue could act as an impetus for developing the kind of expertise needed for future success in the types of complex domains that are likely to be representative of tomorrow. The principles are: (a) *flexibility-focused feedback* (i.e., methods to overcome cognitive rigidity and acquire knowledge flexibly); (b) *concept-case coupling* (i.e., methods that permit learners to experience the different ways in which concepts vary from situation to situation); (c) case-proficiency scaling (i.e., use of mentoring and other scaffolding methods to stretch skill); (d) *tough-case time compression* (i.e., the need to develop a bank of cases, with varying difficulty and complexity, on which to practice adapting); (e) *complexity preservation* (i.e., methods that preserve the functional complexities to be learned and avoid learning oversimplified relationships); (f) active reflection (i.e., methods that help learners become better calibrated in terms of what they know, and in their ability to identify competency boundaries) (see also Havinga et al., this volume). Collectively, these principles were designed to address the need to provide practice at problems that stretch current competency and adaptive capability by promoting the opportunities outlined in Table 1.

Ward et al. (2016, 2018) argued that activities and practices that address the collective objectives outlined in Table 1, through use of the aforementioned principles, is likely to develop the requisite sensemaking and flexecution skills that are integral to any cogent definition of adaptive skill. Measuring developments in adaptive skill and ascertaining the level of expert adaptivity, however, is likely to be a far more challenging task.

Measuring Expertise for Future Work

In the introduction to this Handbook, and in many chapters contained within, several authors discussed the challenges of measuring expertise. Naikar and Brady (this volume) suggest that by traditional standards, in terms of ability to execute an idealized sequence or set of tasks, adaptively skilled performance may be judged as *falling short*—perhaps because it deviates from the original task or results in a *good enough* outcome rather than a *reliably superior* one. However, when viewed through an adaptive lens (i.e., if assessed in terms of ability to engage in flexible action given the circumstances, which may be unfamiliar or unforeseen), performance may be considered effective or even masterful. Given the difficulties of measuring expert performance in many domains, the key question raised in the introductory chapter was: What do you do (and how should one define expertise) when working in the majority of complex domains where performance measurement is particularly challenging or impractical? We add here: What do you do when adaptation is an integral part of performance or a necessary component of expertise, as is likely the case in much future work?

Hoffman, Ford, and Coffey (2000) argued that a proficiency scale for a given domain should be based on multiple methods, and multiple types of method. Specifically Hoffman et al. argued that at least three types of method—referred to as the *three legs of a tripod* (see also Hoffman & Lintern, 2006)—that capture both the breadth (i.e., the variety of relevant experiences) and depth of experience (i.e., the length of those experiences) should be used to scale proficiency. In their research Hoffman et al. used personnel records (e.g., duty assignments), sociometry (or social interaction analysis), and career interviews to gauge skill level and were able to develop a scale that differentiated between levels (expert, journeyman, apprentice) and sub-levels (junior and senior) of expertise (see also Hoffman et al., 2014). Importantly, their method permitted experts to be differentiated from journeymen in terms of their experience and skill at being adaptive.

Hoffman (2018) extended the idea of proficiency scaling based on a three-legged tripod to that of a pentapod, to include: (i) in-depth career interviews to identify breadth and depth of education, training and experience; (ii) professional standards or licensing to identify what it takes for individuals to become an expert at the top of their field; (iii) performance on multiple (rather than one) familiar tasks to identify competence on key problems; (iv) sociometry to identify social networks of who talks to whom about particular problems; and (v) cognitive task analyses to identify mental models of expert knowledge and strategies. Hoffman (2018) argued that one should always use at least two distinct classes of methods in order to obtain evidence that converges on a scale that is appropriate to the given domain, and that permits validation of a proficiency scale. Note that Hoffman's pentapod acknowledges the networked character of expertise and its embeddedness in social environments, as discussed in the beginning of this chapter.

Klein (2018) extended Hoffman's list to include (vi) *within-expert reliability*, (vii) *peer respect*, and (viii) *reflection,* noting that none of these measures alone are fool proof in terms of determining skill level and, like Hoffman, suggesting that multiple measures should be used to measure and scale expertise appropriately. Within-expert reliability refers to the desire for expert performance to be reliably superior, and has been a cornerstone of expert performance measurement for some time (see Ericsson & Ward, 2007).

Peer respect refers to the ability of colleagues to judge one's competence and being able to differentiate those who just talk *a good game* from those who excel in practice (cf. Collins & Evans, this volume). Peer respect should not be confused with interactional expertise since one may have mastered the language of a specialist domain, and hence may be able to interact expertly, but do so in the absence of practical competence or *contributory* expertise (see Collins & Evans, this volume).

Reflection refers to the tendency, and willingness upon request, of an expert to literally reflect on past events and critical decisions made—as a means to identify alternative meanings and courses of actions for any given situation, particularly those that resulted in optimal and sub-optimal outcomes.

Although a number of scale-based measures of adaptive skill have been developed (for a review, see Ward et al., 2016), measurement options *iv* (cognitive task analysis; Hoffman, 2018) and *viii* (reflection; Klein, 2018) above, perhaps, offer a better means to gain insight into the adaptive nature of expertise. These options provide an opportunity to learn whether an expert noticed the need for adaptation as well as how they adapted their understanding, plans, methods, responses, or goals. Others have argued that measures of adaptation that are intimately related to measures of resilience should be incorporated into measurement of expertise, particularly Woods' (1988, 1994; for a review see, Woods, 2017) view of analyzing responses to anomalies at a systems level (see Hoffman & Hancock, 2017). According to this view, successful responding to anomalies requires a model of what is expected, making sense of deviations from those expectations without having to make additional assessment, as well as generating appropriate responses to boot. Such assessments and responses produce more expected and unexpected system responses which require further (2^{nd} order) adaptation (Woods, 2017). As such, we argue that measures that have been proposed to measure resilience capacity may also be used to measure the adaptive nature of expertise at a systems level, as discussed by Hoffman and Hancock (2017).

In future, researchers should attempt to identify which of these measures best differentiates amongst expertise levels in a particular domain, and validate the resulting proficiency scales that are developed with a view to capturing the adaptive nature of expertise within them. In the next section, we provide some additional recommendations for expertise researchers that have emerged from this Handbook.

Future Areas of Research Needed

In addition to a reconsideration of how we view expertise and the kinds of measures we might develop to capture adaptive skill, there are several additional themes that run through a number of chapters and that need to be addressed further in the future. The first is the discussion between structuralism and functionalism. Applied to the field of expertise, this distinction boils down to the question of whether the development of expertise can be explained by a few invariants in human cognition (e.g., the rate with which chunks can be stored in long-term memory), or whether expertise is a matter of tuning to goal-relevant constraints in the environment. The structuralist perspective is highly associated with classical views on expertise, whereas the functionalist perspective is highly associated with ecological and more adaptive views on expertise (i.e., *macrocognition*). We need not be forced to choose between the two, however. As Simon noted, both blades of the scissor, the outer as well as the inner environment, play a role in the development of expertise. People need to learn to discover what the relevant constraints are in the environment they are operating in, and then they need to somehow *learn* to recognize or perhaps *store* these constraints in memory (opinions differ as to whether mental representations are necessary constructs at all in the explanation of expertise, as witnessed by the discussion on situated or enacted/embodied cognition). Future research should take both aspects into consideration, and determine how they jointly operate in the development of expertise.

The outer and the inner environment may also be taken more broadly to include, on the one hand, the physiological environment, and on the other hand the sociological environment. Both our physiology as well as our social environment place constraints on our ability to adapt to task demands. Although both aspects were covered in chapters in this Handbook, future research should look into these aspects in more detail. For instance, with the advent of human enhancement techniques, we may in the not so distant future be able to change our physiological constraints and hence be able to learn faster. Human potential, for example in terms of giftedness, may be unleashed to a fuller extent than is currently possible. Human enhancement may be thought of as intervening in our physiology or even neurology, but it may also consist of social and motivational interventions that target our social learning environment, for instance through games or other forms of simulation or augmented reality. Furthermore, our social learning environments may be extended to include culture and subculture, schools and professional environments. To distinguish the field of expertise studies from already existing fields such as educational psychology and cultural psychology, questions would focus on how to accelerate and sustain expertise through social and motivational interventions (e.g., what is required in terms of social support to sustain deliberate practice on a regular basis over the years; how can we design environments in which scientific creativity can be increased?).

A second recurring theme throughout this Handbook pertains to how we should go about developing expertise and preparing individuals for success in tomorrow's workplace or, indeed, for preserving the expertise that they have at both individual and system levels. Above, we noted that current educational methods might inadequately prepare future experts (see Resnick et al., this volume), we highlighted a different view of learning that might lend itself to supporting the changing demands of complex work (see Klein, 2011), and proposed a set of training principles that could be used to guide the development of adaptive skill (see Ward et al., 2016, 2018). Numerous authors in this Handbook have proposed related ideas about what might be the best methods to develop expertise and adaptive capacity, needed not least because there will always be a gap between how management imagines work and how it is carried out (e.g., Dekker, 2005). We summarize a handful of these here and then offer some thoughts on developing expertise for the future.

Spiro et al. (this volume) argued that "learning and training will likely have to occur in qualitatively different ways to develop the ability to deal adaptively with the resultant increase in novelty that must be routinely confronted." They make several recommendations for developing 21st century skills that will foster adaptivity. For instance, they recommend that we "pay attention to cases in their variegated richness while de-emphasizing the *primacy* of concepts." From their perspective, concepts are still crucial, but they must be tailored to cases rather than vice versa. In addition, they recommended that we "use multiple rather than single conceptual relations (as in schemas, prototypes, analogies, perspectives, etc.); treat cases as wholes with emergent properties so they are greater than the sum of their parts; increase the attunement to difference and decrease the bias toward seeing similarity; expect unpredictability, irregularity, indeterminateness; expect to return to earlier cases in new contexts to bring out facets that were hidden in the earlier context … ; embrace flexibility and openness of knowledge representation over rigidity; stress context-dependency over context-independence; avoid rigidity in understanding, remaining *open* instead, with an appreciation for the sometimes limitless range of uses of knowledge in new combinations, for new purposes, in new situations; rely on situation-adaptive assembly of knowledge and experience rather than retrieval of intact knowledge structures from long-term memory; and so on" (also see Table 1). From our perspective, these features of Cognitive Flexibility Theory and of adaptive skill development are likely to be the cornerstone of learning in future work environments, and the bedrock of training for any future expert.

Another recommendation may be to combine the use of acquisition, refresher, and mobilization training, such that skill relapse and maintenance training are spread throughout the pre- and post- and after-retention intervals (Arthur & Day, this volume). However, while this approach helps mitigate skill decay, traditional overtraining methods could lead to a trade-off between retention and adaptability outcomes: greater retention, theoretically, may lead to more well-grooved behaviors that, potentially, are less conducive to being adapted. One method that may overcome this tradeoff is training that emphasizes learning to spot and respond to anomalies through greater exploration and understanding of the system constraints (Havinga et al., this volume). This is consistent with our previous argument for incorporating measures of responding to anomalies (e.g., Woods, 2017) in any measure of adaptive capability.

Non-training solutions might involve, for instance, the use of regulations and centralized decision making as means to promote adaptability and to help avoid working at cross purposes (Havinga et al., this volume). This recommendation is not too dissimilar to current recommendations by government agencies to support policy officials' decision making. Currently, officials are provided with relevant guidance and tools, such as using *futures* methods (i.e., considering policy options against multiple possible future scenarios). However, Conway and Gore (this volume) note that to foster adaptive decision making such guidance should incorporate the available relevant evidence together with information critical to how this might be instantiated (e.g., as a policy, in practice) in the current context (i.e., political landscape). This should be gathered from a variety of trusted experts and iteratively fused throughout the policy making cycle to ensure that policy is designed with policy maker values, stakeholder responses (e.g., parliament, public opinion), and differential effects in mind and, hence, can facilitate an adaptive response.

Technological solutions to developing and maintaining expertise have also been proposed, such as designing displays and instructions to minimize distracting or irrelevant information and to enhance the most relevant cues (e.g., see Morrow & Azevedo, this volume). One example, provided by Morrow and Azevedo, is a medical phone-based application that could support: (a) physician treatment decisions by providing rapid access to both up-to-date illness/treatment information *and* a flexible decision making workspace that allows this information to be compared with an expert mental model; and (b) patient decisions by providing new illness/treatment information guided by agent-based assessment of patients' current understanding of their illness. However, Moore and Hoffman (this volume) warn against the use of technology that is more of a process control system (i.e., imposes a way of working on the expert) than a decision aid (i.e., informs and supports expert reasoning and decision making). They highlight that such technology, especially the type that integrates and filters data as a means to provide a purported understanding to the user, rather than permitting the (developing) expert to form their own understanding, may be misguided given that the latter drives expert search.

There are several implications of this research for the design of future developmental activities. With some exceptions (for a review, see Hoffman et al., 2014), much of what we know already in terms of how experts learn is based on traditional rather than adaptive measures of expertise or on training derived from studies of non-experts (e.g., undergraduates). Hence, much research is needed to provide the empirical basis for developing experts who are sufficiently prepared to work in future complex work environments. Training that helps build an adaptive mindset (i.e., ability to generalize across cases in a content-independent manner) and adaptive readiness (i.e., skill at situation-adaptive assembly of knowledge and experience), and that does not wholly relinquish this understanding to technology, may be a good contender for developing future expertise (Spiro, personal communication).

Finally, a third recurring theme is the discussion on the generality and specificity of expertise. According to the classic view, expertise is highly domain-specific and does not transfer to novel domains. Other views have challenged this perspective, by emphasizing concepts such as general intelligence, giftedness, or the adaptive nature of expertise (e.g., Lobato, & Siebert, 2002; Ward et al., 2018). Again, expertise may be both general as well as specific, depending on one's view. What seems important here, and as we have discussed at length, is the definition of expertise one proposes. The classic view of expertise seems to reserve the concept for an end state in which there is absolute mastery of the skills and knowledge obtained. In this view, it takes at least ten years of dedicated (deliberate) practice to acquire these cognitive skills. The alternative view seems to view expertise more as a relative concept, in which experts simply possess both more knowledge and skills than non-experts. This view essentially propounds a social perspective on expertise, in which someone is considered an expert if they are considered as such by their peers. Both the classic and the adaptive view on expertise may perhaps be reconciled by taking into account the domains they typically take into consideration: the classic view has mainly studied expertise in scientific, sports or artistic domains, that is, domains where some performance standard can easily be defined and where progress toward some gold standard can be measured. As indicated earlier, the adaptive view on expertise, in contrast, has typically focused on domains where such gold standards are not available, or where practitioners typically do not have the opportunity to spend 10 or more years in the same job—it should be noted that these domains or jobs constitute the majority of real-life situations (think of the frequent job rotations in the military).

This leads to some implications for both practical applications as well as theoretical avenues for research. Over the past 50 years, beginning in the late 1960s, we have made tremendous progress in what we know about expertise. In the first twenty years, the pendulum swung from an emphasis on domain-general characteristics to highly domain-specific characteristics. In the past thirty years, alternative conceptions of expertise have arisen, with more focus on how expertise develops under time-constrained, ill-defined, ambiguous situations with multiple, competing goals. The focus has gradually shifted from explaining expertise in terms of underlying cognitive processes to describing expertise as an adaptation to goal-relevant constraints. However, we may have thrown out the baby with the bathwater in our eagerness to focus on ecologically valid domains. Theoretical depth has given way to theoretical breadth, as an explosion of theoretical concepts occurred with each new description of expertise in a particular domain. This was actually foreseen by Langley and Simon in their concluding chapter of the 1981 book on the "Acquisition of cognitive skills." They stated that as the study of expertise showed increasing domain-specificity as being key to the explanation of expertise, this would make theories of expertise barren and non-generalizable. Instead, they argued, researchers should look at the one remaining constant in the study of expertise, which was the invariance of the learning process itself. Experts in different domains may reach different end points in terms of the contents of their knowledge, what remains constant, or so hypothesized Langley and Simon, is the learning process underlying their course of expertise.

Future researchers should therefore spend more attention to the laws governing the learning process itself, as well as the various ways in which the learning process can be enhanced or accelerated. The important questions of the future are not what expertise is as an end state, but rather how we can turn relative novices into relative experts faster. We may learn some interesting novel laws on how people learn on the way.

A Potential Way Forward

What are the next steps? In this section we discuss two issues: What types of methods and approaches might help us better understand complex behavior, specifically cognitive adaptation to complexity, and the next steps we might take as a field to ensure that our science continues to mature.

In 1973, Allen Newell was challenged with trying to integrate a series of talks on varied aspects of cognition presented by some of the field's best researchers, many of whom were pioneers in the study of expertise. He lamented that it was virtually impossible to synthesize the symposium talks since the associated lines of research were disconnected from one another, and the focus was often framed by disparate polar arguments rather than some common higher goal (e.g., societal need). Despite best intentions to conduct *good science*, the tendency was for this research to be piecemeal rather than systematic, and disaggregated rather than systemic. This state of affairs was partly due to psychology's preferred method of null-hypothesis significance testing, which led to ever-more detailed hypotheses without attempts to develop over-arching frameworks to integrate the results of the experiments carried out. Newell (1973) asserted that the then-current practice resulted in ideological and theoretical differences becoming less clear, and viewpoints never really being combined in a way that moves the field forward. He argued that the end product of such a process was unlikely to be a mature and cumulative science.

As a potential remedy, and with a view to developing what he would later call a "unified theory of cognition" (Newell, 1990), Newell offered three potential approaches for moving the field of cognitive science forward: (i) Develop a complete (rather than partial) processing model that provides a detailed representation of the control structure coupled with equally detailed assumptions about memories and primitive operations, such that what a participant does in completing a task, and how they do it, are fully specified (examples of such complete processing models or *architectures of cognition*, as they would later be called, are SOAR [Newell, 1990] and ACT-R [Anderson, 2007]); (ii) Develop a programmatic approach (using both experimental and theoretical studies) to analyzing a single complete (rather than partial) complex task (i.e., a supertask of tasks or simply, work) such that a model of task behavior can be derived from the program; (iii) Continue in the current vein of conducting disaggregated tasks but build a single processing system (rather than just multiple models) that explains them all and, as a result, their integration. Some of these recommendations have been heeded in the intervening years by researchers in the field of expertise (e.g., see Gobet, this volume; Kirlik & Byrne, this volume; Hoffrage this volume) and, arguably, the research more broadly has become more programmatic. However, the trend has frequently been toward modelling simpler tasks or, when more complex tasks have been modeled (e.g., flying a plane), researchers have focused on partial versions of the task (e.g., taxiing along a runway).

In 1993, David Woods made a similar call to Newell's but placed a greater emphasis on ways to better understand more *complex* phenomena, beyond chess, that occur outside the psychological laboratory. He was particularly interested in suggesting methods that would lend themselves to better understand cognition in natural settings that were "complex, rich, multifaceted… [rather than]…simplified, spartan, single-factor settings" (p. 228). Understanding cognition in natural settings as a form of adaptation to complexity is of relevance to the study of expertise in light of our extensive discussion above on expertise as adaptive skill. Woods suggested several approaches that might be used to advance our understanding of cognitive adaptation to complexity:

a) Establish the mapping between test and target behavioral situations such that the representativeness of the specific test situation adequately reflects the target class of situations of interest and, hence, can be generalized (see Brunswik, 1956; Harris, Foreman, & Eccles, this volume). This is not unlike Ward et al.'s (2018) principle of complexity preservation where the important complexities of target relations are maintained in any test or training;
b) Use cognitive task analysis methods (e.g. verbal reports of thinking, action protocols, data acquisition traces) that externalize signs, which permit inferences to be made about internal processes (see Ward, Wilson, Suss, Woody, & Hoffman, this volume). These are often a precursor to building (e.g., as in verbal or protocol analyses), or are methods of validating, what Newell termed complete processing models;
c) Retrospective analyses of critical incidents where one attempts to reconstruct the mental dynamics that occurred at the time of the event based on participants' interviews and other data (e.g., aviation black box, videos) (see Militello & Shilo, this volume); and
d) Field observations in which the observer immerses themselves in the native environment in order to obtain the point of view of domain practitioners (see Hutchins, 1995; Yardley, Mattick, & Dornan, this volume).

Woods (1993) argued that the methods he proposed—especially when studying experts—permitted complex worlds to be examined directly in a way that produces results relevant to the local problem while, simultaneously, contributing results to the generic research base on human cognition. He referred to this perspective as a complementarity assumption where one could, as

Stokes (1999) later asserted, pursue both a quest for understanding and a quest for utility because these dimensions are orthogonal rather than represent ends of a single basic-applied continuum. This view promotes the use of specific behavioral contexts, which are representative of a particular class of behavior or environment, as field laboratories for examining these behaviors. Woods (2003) later referred to this approach as “staged world observation,” which was more geared to verification (rather than discovery) than natural history methods (pure observation of what naturally occurs), but which shaped the conditions of observation to a lesser extent than what he referred to as “Spartan Lab” experiments (experimenter-created problem situations with low authenticity). In staged world studies, the authenticity depends on the investigators’ ability to design scenarios and stage situations of interest: the scenarios need to be recognized as valid by domain practitioners (or experts in the field). Many researchers have followed this guidance and have studied experts (and expert-novice differences) in scaled world simulations using process tracing measures (for instance, Sarter & Woods, 1995, in their studies on mode confusion in the cockpit or Schaafstal, Schraagen, & Van Berlo, 2000, in their studies on electronics troubleshooting).

Other researchers have produced cognitive models of expertise (e.g., see Burns, this volume; Hutton, this volume; Kirlik & Byrne, this volume; Matthews, Wohleber & Lin, this volume; Ross & Phillips, this volume; Salmon, Stanton, Walker, & Read, this volume), albeit of a qualitatively different type to that proposed by Newell (1973). These models are illustrative of the field of the maturation of the field of expertise studies (i.e., a focus on descriptive and product rather than process models) and, importantly, have played a vital role in improving our understanding of expertise in complex work settings, and in helping develop complementary theoretical advances that are more broadly applicable.

Despite this substantial contribution, there has been a tendency for some researchers to value only one type of research context or method. Paraphrasing Gigerenzer (2004), Ward, Belling, Petushek, and Ehrliger (2017) acknowledged this tension by highlighting that "disdain has routinely been expressed by a diverse range of scientists for those 'in the other camp,' whose position, purpose, and methods have been described by those holding contrary views as having little scientific or societal value" (p. 18). Woods (1993) described such views as being *destructive,* and suggested that views that upheld one methodological strategy as having more privileged access to fundamental results than another does little to build a complete understanding of expertise. As we argued in Ward et al. (this volume), no single method or context should hold precedence if we are ever to develop a fuller understanding of expertise (see Hoffman, 2018). Despite the propensity of some to elevate one methodological approach over another (e.g., Banaji & Crowder, 1989), the complementarity between naturalism and experimentalism has long been appreciated (e.g., by Darwin) and has recently been reiterated (see Klein et al., 2003; Woods, 2003). Accordingly, we argue that both Newell's and Woods' respective recommendations provide useful guidance for how we should continue to advance the study of expertise. The complementarity assumption still holds today and is likely to prove useful guidance tomorrow.

Although some in the field of expertise studies have frequently advocated that researchers adopt a specific methodological approach or theoretical framework, few communities of practice have actually taken stock of the progress they have made towards their research being cumulative, or their community producing mature science, at least not in the reflective manner employed by Newell (1973). In a rare exception, Klein (2017) reviewed the progress of the Naturalistic Decision Making and Macrocognition Community of Practice and made some recommendations for where we go next. In terms of progress, Klein stated that the communities had moved from being research-oriented to being more practice-oriented, providing services to customers in very diverse domains (e.g., Crichton & Moffat, this volume; Moon, this volume; Roth, Naweed & Multer, this volume; Wiggins, Auton, & Taylor, this volume).

In terms of recommendations, Klein highlighted the fact that many of the models developed to this point were descriptive rather than predictive, and suggested that more effort was needed to translate these models into reliably effective interventions and tools. Second, he pointed out that much of the research had been conducted at the tactical level (e.g., examining boots-on-the-ground decision making) rather than at the operational (e.g., command) or strategic (e.g., policy) levels, and there was little guidance to help researchers and practitioners translate between levels, or empirical research available to corroborate that what worked at one level could be successfully applied at another. Last, it had become popular for researchers from communities beyond the expertise super community to demonize experts and to devalue their contribution to society (see Klein et al., this volume). Hence, all of the expertise communities of practice needed not just to advance the field but to do so in the face of external antagonism.

Klein's reflection indicates that much progress has been made in the study of expertise. However, Klein noted many challenges still lie ahead that are relevant to most if not all expertise communities of practice: There is a need to better understand and delineate the cognitive and social processes underlying the many micro- and macro-cognitive functions that have been described to date. There is a requirement to address the needs of strategic decision makers at an organizational and policy level. There is a need for a clear public message about what the science of expertise can offer – especially in light of the war on experts – and to increase public and professional awareness and global reach of this super community. Last, there is a need to further develop instructional opportunities to better prepare scientists and practitioners to use the types of the methods discussed in this Handbook. These are just some of the challenges for the budding expertise researcher. Our hope is that this Handbook can act as a catalyst to help future researchers address some of these challenges as well as address the current issues and future directions highlighted above and throughout the Handbook.

References

Anderson, J.R. (2007). *How can the mind occur in the physical universe*? New York, NY: Oxford University Press.

Banaji, M. R., & Crowder, R. G. (1989). The bankruptcy of everyday memory. *American Psychologist, 44*(9), 1185-1193.

Dekker, S. W. A. (2005). *Ten questions about human error: A new view of human factors and system safety.* London, UK: Lawrence Erlbaum Associates.

Eiser, J. R., Stafford, T., Henneberry, J., & Catney, P. (2009). "Trust me, I'm a scientist (not a developer)": Perceived expertise and motives as predictors of trust in assessment of risk from contaminated land. *Risk Analysis, 29*(2), 288-297.

Ericsson, K. A., & Ward, P. (2007). Capturing the naturally-occurring superior performance of experts in the laboratory: Toward a science of expert and exceptional performance. *Current Directions in Psychological Science, 16*(6), 346-350.

Fadde, P. J., & Klein, G. (2010). Deliberate performance: Accelerating expertise in natural settings. *Performance Improvement*, *49*(9), 5–14.

Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics, 33*, 587–606.

Hatano, G., & Inagaki, K. (1984). Two courses of expertise. *Research and Clinical Center for Child Development Annual Report*, *6*, 27–36. http://hdl.handle.net/2115/25206

Hatano, G., & Inagaki, K. (1986). Two courses of expertise. In H. Stevenson, H. Azuma, & K. Hakuta (Eds.), *Child development and education in Japan* (pp. 262–272). New York, NY: W. H. Freeman (Chapter 17).

Hoffman, R. R. (1998). How can expertise be defined? Implications of research from cognitive psychology. In R. Williams, W. Faulkner, & J. Fleck (Eds.), *Exploring expertise* (pp. 81–

100). New York: Macmillan (Chapter 3).

Hoffman, R. R. (2018). *Pentapod.* Report, Institute for Human and Machine Cognition, Pensacola, FL.

Hoffman, R. R., Ford, K. M., & Coffey, J. W. (2000). *The Handbook of human-centered computing*. Report, Institute for Human and Machine Cognition, Pensacola, FL.

Hoffman, R. R. & Hancock, P. A. (2017). Measuring resilience. *Human Factors, 59*(4), 564-581. Doi: 10.1177/0018720816686248.

Hoffman, R. R., LaDue, D., Mogil, H. M., Roebber, P., & Trafton, J. G. (2017). *Minding the weather: How expert forecasters think*. Cambridge, MA: MIT Press.

Hoffman, R. R., & Lintern, G. (2006). Eliciting and representing the knowledge of experts. In K. A., Ericsson, N. Charness, P. Feltovich, & R. Hoffman, (Eds.). *Cambridge Handbook of expertise and expert performance* (pp. 203-222). New York: Cambridge University Press.

Hoffman, R. R., Ward, P., DiBello, L., Feltovich, P. J., Fiore, S. M., & Andrews, D. (2014). *Accelerated expertise: Training for high proficiency in a complex world*. Boca Raton, FL: Taylor and Francis/CRC Press.

Hutton, R., Ward, P., & Turner, P. (2018). *Testing the adaptive performance principles: Design of an evaluation study for Intermediate Command & Staff Course (Land) (Final Technical Report v2).* O-DHCSTC-I309549_PT_T3_234/008. Yeovil, UK: BAE Systems.

Hutchins, E. (1995). *Cognition in the wild.* Cambridge, MA: MIT Press.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist, 64*(6), 515-26.

Klein, G. (2011). *Streetlights & shadows: Searching for the keys to adaptive decision making* (pp. 1–337). Cambridge, MA: MIT Press.

Klein, G. (2018). How can we identify experts: Seven criteria for deciding who is really credible. *Psychology Today* (Sep 01, 2018). Retrieved from https://www.psychologytoday.com/us/blog/seeing-what-others-dont/201809/how-can-we-identify-the-experts

Klein, G., & Baxter, H. C. (2006). Cognitive transformation theory: Contrasting cognitive and behavioral learning. In *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)* (pp. 1–9). Arlington, VA: NTSA.

Klein, G., & Baxter, H. C. (2009). Cognitive transformation theory: Contrasting cognitive and behavioral learning. In D. Schmorrow, J. Cohn, & D. Nicholson (Eds.), *The PSI Handbook of virtual environments for training and education: Developments for the military and beyond. Volume I: Learning, requirements and metrics* (pp. 50–65). Westport, CT: Praeger Security International (Chapter 3).

Klein, G. A., & Hoffman, R. R. (1993). Seeing the invisible: Perceptual/cognitive aspects of expertise. In M. Rabinowitz (Ed.), *Cognitive science foundations of instruction* (pp. 203-226). Mahwah, NJ: Lawrence Erlbaum Associates.

Klein, G., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., & Hollnagel, E. (2003). Macrocognition. *IEEE Intelligent Systems, 18*(3), 81-85.

Kofman, A. (2018). "Bruno Latour, the post-truth philosopher, mounts a defense of science." The New York Times Magazine, October 25.

Langley, P., & Simon, H. A. (1981). The central role of learning in cognition. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Lawrence Erlbaum.

Latour, B., & Woolgar, S. (1979). *Laboratory life: The construction of scientific facts*. Princeton, NJ: Princeton University Press.

Latour, B. (2018). *Down to earth: Politics in the new climatic regi*me. Cambridge, UK: Polity Press.

Lobato, J., & Siebert, D. (2002). Quantitative reasoning in a reconceived view of transfer. *Journal of Mathematical Behavior, 21*, 87-116.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W.G. Chase (Ed.), Visual information processing (pp. 283-310). New York, NY: Academic Press.

Newell, A. (1990). Unified theories of cognition. Cambridge, MA: Harvard University Press.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering.* New York, NY: North-Holland/Elsevier.

Rasmussen, R., Pejtersen, A. M., & Goodstein, L. P. (1994). Cognitive systems engineering. New York, NY: Wiley & Sons.

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors, 37*(1), 5-19.

Schaafstal, A. M., Schraagen, J. M. C., & Van Berlo, M. (2000). Cognitive task analysis and innovation of training: The case of structured troubleshooting. *Human Factors, 42*(1), 75-86.

Shanteau, J. (1992). Competence in experts: The role of task characteristics. *Organizational Behavior and Human Decision Processes, 53*, 252-266.

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1992). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. In T. M. Duffy, & D. H. Jonassen (Eds.), *Constructivism and the technology of instruction: A conversation* (pp. 57–76). Hills- dale, NJ: Lawerence Erlbaum Associates (Chapter 5).

Stokes, D. E. (1997). Pasteur's quadrant: *Basic science and technological innovation.* Washington, DC: Brookings Institution Press. ISBN 9780815781776.

Ward, P., Belling, P., Petushek, P., & Ehrlinger, J. (2017). Does talent exist? A re-evaluation of the nature-nurture debate. In J. Baker, S. Cobley, J. Schorer, & N. Wattie (Eds.), *Routledge handbook of talent identification and development in sport* (pp. 19-34). London, UK: Routledge.

Ward, P., Hutton, R., Hoffman, R. R., Gore, J., Anderson, T., & Leggatt, A. (2016). Developing skilled adaptive performance: A scoping study (Final Technical Report v3). In *O-DHCSTC I2 T T2 077 002*. Yeovil, UK: BAE Systems.

Ward, P., Gore, J., Hutton, R., Conway, G., & Hoffman, R. (2018). Adaptive skill as the *conditio sine qua non* of expertise. *Journal of Applied Research in Memory and Cognition, 7*(1), 35-50.

Woods, D. D. (1988). Coping with complexity: The psychology of human behavior in complex systems. In L. P. Goodstein, H. B. Andersen, & S. E. Olsen (Eds) *Mental models, tasks and errors* (pp. 128-148). London: Taylor and Francis.

Woods, D. D. (1993). Process-tracing methods for the study of cognition outside the experimental psychological laboratory. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), Decision making in action: Models and methods (pp. 228-251). Norwood, NJ: Ablex Publishers

Woods, D. D. (1994). Cognitive demands and activities in dynamic fault management: Abduction and disturbance management. In N. Stanton, (Ed.), *Human factors of alarm design* (pp. 63-92). London: Taylor and Francis.

Woods, D. D. (2003). Discovering how distributed cognitive systems work. In E. Hollnagel (ed.), *Handbook of cognitive task design* (pp. 37-53). Mahwah, NJ: Lawrence Erlbaum Associates.

Woods, D. D. (2017). Reflections on the origins of cognitive systems engineering. In P. Smith and R. R. Hoffman (Eds.), *Cognitive systems engineering: The future for a changing world.* Boca Raton, FL: Taylor and Francis.

Table 1: Collective objectives of Ward et al.'s (2016, 2018) Adaptive Skill Training Principles (adapted from Ward, Hutton, Hoffman, Gore, Anderson, & Leggatt, 2016; pp. 20-21).