Advances in Cognitive Systems 8 (2019) 13–32 Submitted 5/2019; published 12/2019

The Understanding Problem in Cognitive Science

Alexander R. Hough HOUGH.15@WRIGHT.EDU
ORISE at Air Force Research Laboratory, Wright-Patterson Air Force Base, OH 45433 USA
Wright State University, Psychology Department, Dayton, OH 45435 USA

Kevin A. Gluck KEVIN.GLUCK@US.AF.MIL
Cognitive Models and Agents Branch, Air Force Research Laboratory, Wright-Patterson Air Force Base, OH 45433 USA

Abstract

Understanding is at the core of higher-level information processing and has a long history in the cognitive sciences. It is often described as a complex phenomenon with many dimensions, which makes it difficult to define with precision. Many researchers have noted that understanding is often ill defined, indirectly addressed, or avoided altogether. This is particularly disappointing considering that understanding has been a topic of interest since the ancient Greeks. In order to address this problem with our understanding of understanding, we reviewed literature from philosophy, psychology, education, neuroscience, and computer science. Here we summarize insights from that review, focusing on similarities and differences across those domains, as well as implications for the nature, measurement, and modeling of understanding.

1. Perspectives Across Disciplines

Understanding has been discussed and debated for millennia, yet remains rather poorly understood. There have been efforts in cognitive science to describe, measure, and model human understanding. These efforts are often incomplete, domain specific, or at too high a level to describe the underlying processes with precision. This state of affairs is disappointing for many of us working in the cognitive sciences, and the difficulties in defining and implementing understanding are often acknowledged (e.g., Gluck & Laird, 2019a, 2019b; Hayes & Simon, 1974; Langley, Laird, & Rogers, 2009; Pritchard, 2009; Rogers, 2008; MacLellan, Harpstead, Marinier, & Koedinger, 2018; Minsky, 1975; Moore & Newell, 1974; Schank & Abelson, 1977; Simon, 1980, 1986; Simon & Hayes, 1976; Smith & Siegel, 2004). Here we report a multi-disciplinary literature review to address the lack of consensus and clarity in the literature. Our goal was to find common ground in order to facilitate cross-domain collaboration and progress. We discuss similarities and differences in conceptualizations, measurement, and modeling of understanding, and conclude by providing a working definition of understanding with some suggestions for future measurement and modeling.

Although the topic is under-researched in general, several fields of study are relevant for improving our scientific understanding of understanding. Here we review conceptualizations from philosophy, psychology, education, neuroscience, and computer science. We use this review to identify eight common features from across these disciplines. These features may serve as the basis for an emerging consensus regarding the nature of understanding.

Philosophers emphasize the role of schemas and language to understand the world (Kant, 1958; Wittgenstein, 1961). Relations and context appear more important than knowledge (Abelson, 1980; Austin, 1962; Skyrms, 1980; Woodward, 2003; Zagzebski, 2001) and some consider knowledge to be separate from understanding (Grimm, 2006; Searle, 1980). Skyrms (1980) and Woodward (2003) suggest understanding should let one estimate or predict how things could be different if one variable in the situation were to change. This suggestion could involve systematic manipulation and scientific predictions. However, performance may not be sufficient to measure understanding, as one could perform well or infer correctly by luck with faulty knowledge (Gettier, 1963; Grimm, 2006; Pritchard, 2009, 2016; Rorty, 1979; Searle, 1980). In philosophy, understanding requires valid inferences, appropriate performance, and valid or verified knowledge.

Psychological scientists emphasize that understanding is a matter of degree and occurs incrementally. This process often starts at a superficial level and moves towards deeper, more meaningful concepts and relations, with analogical reasoning often serving as an important component process (Gentner, 1983; Gick & Holyoak, 1983; Katona, 1940; Kintsch & Greeno, 1985; Simon, 1980). Understanding requires domain knowledge and higher-order cognitive skills that allow for active organization, reorganization, flexible utilization of knowledge, and ability to generate or construct concepts or procedures (Anderson & Schunn, 2000; Bransford, & Johnson, 1972; Greeno, Riley, & Gelman, 1984; Greeno, 1977; Schank, 1980; Simon, 1986). In order to understand, one needs to acquire a knowledge base of organized information with rich sets of relations that can generalize to novel situations (Forbus & Gentner, 1997; Hofstadter, 2001; Scholz, 2018). Assessing one's level of understanding and knowledge through metacognitive monitoring is important to identify faulty knowledge or to trigger search for additional information (Butcher & Sumner, 2011).

Education scientists emphasize the interaction between higher-order cognitive skills and knowledge. These skills are important for organizing or structuring incoming information (Perkins, 1998; Perkins & Simmons, 1988), framing problems (Schoenfeld, 1992), identifying understanding gaps (Mayer, 1998), and utilizing context to reduce ambiguity (Brenner et al., 1997). Wertheimer's (1959) *Productive Thinking* proposed that teaching with multiple perspectives and relations is more effective than unidimensional methods. This allows for more flexible use of knowledge, such as the ability to adapt to novel math problems with only slight modification of learned procedures (Greeno et al., 1984). This type of learning is effective, but often requires considerable time, examples, and effort, which may not be practical when there are several learners, time constraints, or low probability of knowledge transfer (Anderson, Reder, & Simon, 1997). In educational settings, performance often serves as an indication of understanding (e.g., Perkins, 1998). However, as we will see later, using performance as a measure is not straightforward and can be misleading.

Cognitive neuroscientists provide evidence that understanding is a complex distributed process and emphasize the interactions among brain areas. Understanding-related processes appear to be neuro-functionally distributed (Bunge, 2004; Bunge, Wendelken, Badre, & Wagner, 2005), and there is evidence for the importance of semantic relatedness (Whitaker, Vendetti, Wendelken, & Bunge, 2018) and meaningful context (Prat, Mason, & Just, 2012). Neuroimaging data suggest knowledge can be dissociated from understanding to a degree (Bunge 2004; Green, Vendetti, Wendelken, & Bunge, 2006), which corresponds with philosophical arguments (e.g., Grimm, 2006;

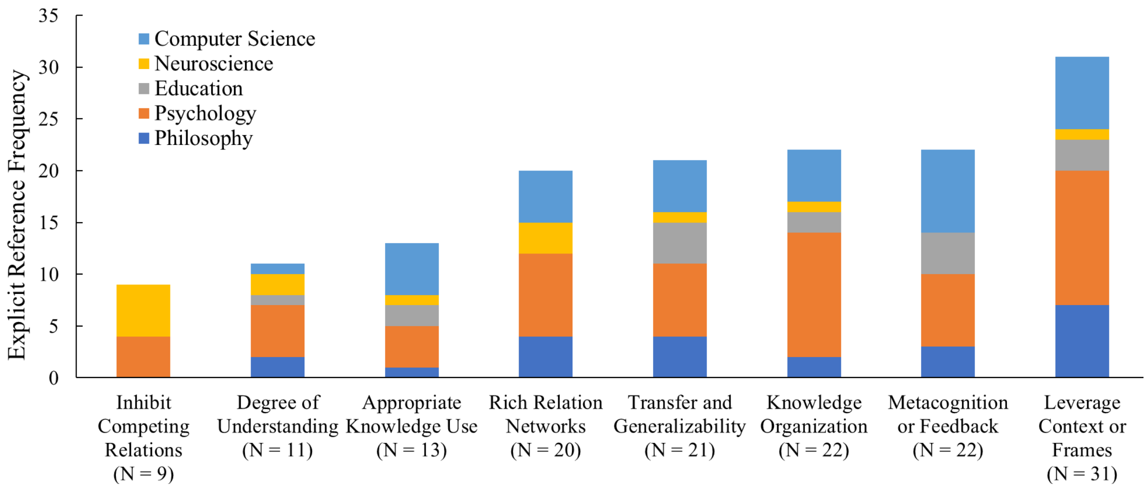

Figure 1. Overall and domain frequencies of explicit reference to the common features of understanding.

Searle, 1980) noted earlier. Understanding is associated with increased efficiency, including controlled semantic retrieval (Whitaker et al., 2018) and inhibition of competing relations (Krawczyk, 2010), measured by changes in brain activation (Ischebeck, Zamarian, Schocke, & Delazer, 2009).

Computer scientists compare understanding-related processing in humans and computers. They note that understanding involves (1) performance (Lindsay, 1963), (2) incremental steps (Moore & Newell, 1974), (3) efficient search and use of indexed information (Simon, 1980; Simon & Gilmartin, 1973), (4) networks of relations and the ability to perform in novel situations (Eliasmith & Thagard, 2001; Forbus & Gentner, 1997), and (5) information about the environment (Schank & Abelson, 1977; Simon, 1980). Some argue for unorganized encoding so analogies can be actively constructed upon retrieval (Veloso & Carbonell, 1993), while others suggest organized encoding for efficient identification and processing of relations (Simon, 1980). Perhaps a hierarchically-organized, nearly-decomposable, system could allow a combination of these possibilities (Greeno et al. 1984; Simon, 1996; Schoenfeld, 1992). Varied or distributed representations may allow for flexibility, generativity, and creative use of knowledge (Hummel & Holyoak, 2005; Langley et al., 2009; Shanahan, 2006). As in psychology, metacognition and feedback are seen as important for identifying faulty or missing knowledge (Forbus & Hinrichs, 2006; Kirk & Laird, 2014; Marshall & Hofstadter, 1997; Thomson et al., 2012).

Eight common features of understanding emerge from these diverse perspectives (see Figure 1 above). Understanding involves (1) degrees of understanding, (2) appropriate use of knowledge, (3) organized knowledge structures, (4) generalization and transfer of knowledge, (5) rich networks of relations, (6) leveraging context or frames, (7) inhibition of competing relations or representations, and (8) metacognition or feedback.

These common features relate to an assortment of representations and processes associated with higher-level cognition. One salient example is the role of analogical reasoning, which involves relational correspondence between a new representation or problem and a previously encountered one (e.g., Gentner, 1983). These representations are organized in a network of relations that progress from surface-level similarity to more abstract, general, and transferable relations. Similarly, Gick and Holyoak (1983) emphasized the construction of schemas or multiple converged representations that reflect deep structural relations, are general, and not bound to a specific context. During their experiments, dissimilar analogies increased the quality of schemas by increasing focus on deeper relations. Analogical reasoning involves knowledge manipulation skills and provides an important example of a higher-level cognitive process that contributes to understanding.

In addition to commonalities with analogical reasoning, the eight features of understanding also emphasize the interaction of multiple components or systems and the benefits of metacognition or feedback to inform progress, identify missing or faulty knowledge, and help calibrate mental representations to the external environment. The role of inhibitory processes in promoting understanding was mentioned primarily in the psychology and neuroscience literatures. It is included in Figure 1 to highlight the possible implications for relevant mechanisms in computational systems intending to explain and predict degrees of understanding.

2. Measurement of Understanding

Degree of understanding is typically measured via performance (Langley et al., 2009; Perkins, 1998; Simon, 1980). Perkins (1998) argues that performance is a criterion of understanding that involves explanation behind an action or solution, building an argument, and going beyond demonstrations of rote memory. As in critical thinking, one should be able to explain, justify, extrapolate, relate, and apply in ways that go beyond knowledge and skill (Halpern, 1998, 2014). Comprehension questions after reading text are good examples. They require generativity and construction, rather than merely reproducing information or knowledge.

Perkins (1998) suggests that understanding involves building a "repertoire of complex performances" (p. 52). This is similar to Greeno et al.'s (1984) generative view of competence. Here, one can perform in a novel situation or with changing task conditions due to the complex interaction of different kinds of knowledge, flexibility, and/or slight modification of procedures. Langley et al. (2009) suggest exposing an agent to novel situations, tasks, and objects that are not directly related to previous experience, then measuring performance to assess understanding. These interpretations imply that the extent of understanding can be measured by how far a person or system can go beyond the explicit information that was given.

Performance is a commonly used indicator of understanding, but there are certainly issues with using it for this purpose. An individual could perform satisfactorily without actually understanding (Greeno et al. 1984; Simon, 1980), and several cases have been documented (Ericsson & Harris, 1990; Ericsson & Oliver, 1989). Perkins (1998) recommends measuring performance on diverse tasks with varying requirements to assess flexibility, as an indicant of the degree of understanding achieved. Including some additional measures such as verbal protocols, eye tracking, or brain imaging could help validate these behavioral-based measurements. This type of converging measures methodology has proven informative in categorization (Gluck, Staszewski, Richman, Simon, & Delahanty, 2001) and decision making (Walsh, Collins, & Gluck, 2015; Walsh & Gluck, 2016), and may also be useful for more complete measurement of degree of understanding.

2.1 Higher-Order Cognitive Skills

Measurements of understanding typically involve tests of domain knowledge. However, some higher-order skills are needed, as knowledge itself is not sufficient for understanding. Some examples are the ability to engage in metacognition (Mayer, 1998), organize information (Perkins, 1998; Perkins & Simmons, 1988), construct relevant representations (Brenner et al., 1997; Schoenfeld, 1992), and recognize when previous knowledge can be used in a novel situation (Clarke, Cheeseman, Roche, & van der Schans, 2014). These findings led Butcher and Sumner (2011) to suggest a sensemaking paradox where higher-order skills are required to learn complex topics, but domain knowledge is needed to apply them. These skills are similar to critical thinking, which Halpern (2014, p. 6) described as "applying knowledge in real-world settings, analyzing and solving problems, connecting choices to actions, and being able to innovate and be creative".

Halpern (1998) suggested training in critical thinking and metacognitive monitoring can enhance understanding and transfer across domains, which was later supported by a meta-analysis (Abrami et al., 2008). The meta-analysis also identified several measures of critical thinking: the Watson-Glaser Critical Thinking Appraisal (Watson & Glaser, 1980), the Cornell Critical Thinking Test (Ennis & Millman, 1985), and the California Critical Thinking Skills Test (Facione, 1990). A more recent measure of high-order cognition, the Halpern Critical Thinking Assessment (Halpern, 2010), includes decision making, problem solving, hypothesis testing, argument analysis, likelihoods and uncertainties, and verbal reasoning. Considering the comprehensiveness of these tests, it may be beneficial to measure domain understanding by using this methodology and including the desired subject matter.

2.2 Measuring Understanding in Models

Performance-based measurements are the most popular technique to assess understanding in humans, but those measurement processes would need to be modified to assess model understanding. Cognitive models tend to be restricted to specific tasks, domains, or cognitive phenomena related to laboratory experiments (Laird et al., 2017; Langley et al., 2009; MacLellan et al., 2018; Simon, 1980; Thomson et al., 2012). This restriction complicates assessing generalizability and higher-order skills. Addressing understanding may require a shift to more large scale or gestalt-like phenomena (Ericsson, Polson, & Wertheimer, 1982; Schierwagen, 2009; Wertheimer, 1985), which would require more complex cognitive models that could later be broken down into components for precise empirical testing (e.g., Ball, 2008).

3. Understanding-Related Processes

Multiple perspectives during learning can foster productive thinking and interest (Wertheimer, 1959), while narrow approaches may decrease interest and foster "re-productive thinking" or rigid knowledge application (Cunningham & MacGregor, 2014, p. 47). However, balancing time and effort costs with the benefits of transfer and practical importance of understanding the subject are also important (Anderson, Greeno, Reder, & Simon, 2000; Anderson et al., 1997; Greeno, 1997).

3.1 Organization of Memory

Although organization of memory relates to performance and degree of understanding (Mannes & Kintsch, 1987), it requires time and effort, and it is subject to a "preparation-deliberation tradeoff" (Newell, 1994). Individuals may not put forth the effort required to understand if it does not serve a future purpose or is not required to solve the task at hand (e.g., Gick & Holyoak, 1983). In addition, people often do not have an accurate assessment of their own understanding. Comprehension may be assumed unless there is some kind of error signal, such as an unfamiliar word or concept, slowing down of processing, or a need to exert more effort (Glenberg, Wilkinson, & Epstein, 1982; McNamara, Kintsch, Songer, & Kintsch, 1996). As seen with experts, this process requires less effort as knowledge increases, becomes more organized, and featural correspondences are recognized automatically (Chase & Simon, 1973; Klein, 1998; Simon & Gilmartin, 1973). However, it may be impossible for understanding to become completely automatic or unconscious because one must know the how, the why, or both (Kohler, 1929; Prichard 2016; Scholz, 2018).

Experts provide insight into understanding, by demonstrating mastery in their domains. Expert chess players use a great deal of knowledge and prior experience in short periods of time by using previously organized and indexed relation-based information, and leveraging both short-term and long-term memory structures (Chase & Simon, 1973; de Groot, 1965; Gobet & Simon, 1996; Simon & Gilmartin, 1973). Experts recognize correspondence between existing representations and a current situation, recall the representation that most closely matches the situation, and then act quickly and efficiently (Klein, 1998). Simon (1986) suggested the recognition of relations and indexed knowledge to be at the core of understanding.

As the sensemaking paradox (Butcher & Sumner, 2011) suggested, domain knowledge is needed to organize or chunk information by meaning or relevance, to recognize features of situations, and to retrieve the appropriate representations (Chase & Simon, 1973; Hofstadter, 2001). For instance, novices often attempt to solve physics problems by working backwards, whereas experts retrieve formulas and develop a plan while reading the problem (Larkin, McDermott, Simon, & Simon, 1980; Simon & Simon, 1978). Experts possess more knowledge, but their organization of knowledge and ability to represent problems by focusing on the concepts and deep structure underlies their skills (Chi, Glaser, & Rees, 1982; Ericsson & Lemann, 1996). A novice chess player can reproduce expert-level performance after spending approximately 50 hours memorizing chess positions. However, they tend to focus on salient chess pieces rather than having a deeper understanding of relations among them (Ericsson & Harris, 1990; Ericsson & Oliver, 1989).

3.2 Analogical and Relational Reasoning

Studying children provides insight into the development of higher-order cognitive skills and their interaction with relevant knowledge. They often have the necessary knowledge, but less control over semantic retrieval and ability to inhibit competing representations. Whitaker et al. (2018) found that children show activation of the same brain network as adults, but activation associated with fine-tuned control of sematic retrieval increases with age. Children also make errors because of faulty knowledge utilization (Greeno et al., 1984). Even when the relevant relations are known, they may still make errors, but this is more likely before a proposed relational shift from surface to deep features (Richland, Morrison, & Holyoak, 2006).

Working memory constraints and cognitive load are also issues for understanding. After controlling for age, Simms, Frausel, and Richland (2018) found working memory capacity was a significant predictor of children's analogical reasoning ability. These results highlighted the importance of schema-like structures (Sweller, van Merrienboer, & Paas, 1998), consolidating representations (Johnson-Laird, 2013), and the ability to inhibit less relevant relations (Richland et al., 2006; Zelazo, Frye, & Rapus, 1996). Individuals with higher working memory capacity or processing efficiency are able to maintain more semantically relevant information during a task (Prat et al., 2012), which appears to be important for analogical reasoning (Simms et al., 2018).

In adults, Gentner's (1983) structure mapping theory suggests progression from surface similarity to analogies, then to general laws. Gick and Holyoak (1983) suggest a similar progression of abstraction from a core idea or mapped identity to a more refined convergence schema that is better able to transfer to novel problems. However, in their study, participants had difficulty recognizing deep relations between existing and novel problems without a hint. Interestingly, Prat et al. (2012) found that giving hints based on superficial relations did not improve performance, which emphasizes the importance of deep relations in analogies. People may not focus on processing that leads to a problem solution. However, when given instructions to do so, this information is acquired through metacognitive monitoring (Berardi-Coletta, Buyer, Dominowski, & Rellinger, 1995).

3.3 Persistence, Creativity, and Reframing

Perkins (1998) suggested understanding is an active process that incrementally builds with experience, requires one to be challenged, and involves conflicts with existing understanding. Wertheimer (1959) stressed the importance of restructuring or reorganizing problem representations after open-minded, creative exploration, which involves flexible use of knowledge. Two examples are Maier's (1931) string problem and Duncker's candle problem (1945). Both require the use of objects in creative ways. These appear to be solved by insight and may involve a new understanding or realization of task features that were not thought of originally (Pols, 2002). Subsequently, the problem can be restructured and solved. Insight appears to be less deliberate and involves realization of an important cue that directs attention towards a relevant analogy (Langley & Jones, 1988).

3.4 Metacognition

An individual or agent may stop information search or deliberation if understanding is believed to have been achieved. Metacognition appears to be important for self evaluation of understanding and resulting effort allocation on tasks like the Cognitive Reflection Test (CRT; Frederick, 2005) and Wason card selection (Wason, 1966). Those with better metacognitive monitoring were more likely to respond correctly during the CRT (Mata, Ferreira, & Sherman 2012) and Wason (Thompson, Evans, & Campbell, 2013). In several experiments, Thompson, Prowse Turner, and Pennycook (2011) demonstrated that a subjective confidence judgment, the Feeling of Rightness, influences whether an individual should give an answer or engage in further processing. The performance-based CRT and subjective Feeling of Rightness both appear to tap into metacognitive monitoring and perceived congruence between one's response and task structure. This conflict or mismatch has also been measured using response times (De Neys & Glumicic, 2008; Pennycook, Fugelsang, & Koehler, 2015) and revealed individual differences (Mata, Ferreira, Voss, & Kollei, 2017).

Consensus suggests a critical function of metacognitive monitoring is to detect conflict or mismatch between an environment and a strategy, type of processing, or expectation. This is influenced by the predictability of the environment, presence of diagnostic cues, and whether goals are reachable (Dannenhauer, Cox, & Muñoz-Avila, 2018; Evans & Stanovich, 2013; Johnson, Roberts, Apker, & Aha, 2016; Pennycook et al., 2015; Stanovich, 2018; Vattam, Klenk, Molineaux, & Aha, 2013). Conflict detection triggers the need for a different approach to problem solving or task completion (Butcher & Sumner, 2011; Dannenahuer et al., 2018; Pennycook, 2017; Stanovich, 2018).

Some suggest that metacognitive ability is more closely related to actual intelligence than traditional measures, because it includes motivation and ability (Frederick, 2005; Toplak, West, & Stanovich, 2011). Overall, self feedback through metacognition helps evaluate the degree of understanding by identifying faulty knowledge representations or discrepancies between current and required understanding, and may signal the need to seek more information or further processing.

3.5 Goal Direction and Purpose

The development of understanding is an intentional, goal-directed, purposeful act. It may be more efficiently achieved when that purpose or goal is explicitly known. However, in real-world situations, the precise goal is not always explicit or salient. The ability to generate, use, and reason about goals to direct behavior is often considered an indicator of intelligence (Newell & Simon, 1972; Schank & Abelson, 1977). These skills are important for autonomy and involve interacting competencies, such as reasoning and problem solving (Johnson et al., 2016; Vattam et al., 2013).

Researchers focused on internal goals (i.e., goal-driven autonomy) have developed agents that autonomously formulate goals or plans, based on situation awareness (Dannenhauer et al., 2018). This approach leverages discrepancies between expectations and the environment, and when detected, they are addressed by modifying goals (Cox, 2007; Klenk, Molineaux, & Aha, 2013; Muñoz-Avila & Cox, 2008), learning (Ram & Leake, 1995), and reasoning about goals (Aha, 2018; Roberts, Borrajo, Cox, & Yorke-Smith, 2018). When goals are generated, their selection can be domain specific (Shapiro, Sardina, Thangarajah, Cavedon, & Padgham, 2012), domain independent (Wilson, Molineaux, & Aha, 2013), or based on priorities (Young & Hawes, 2012).

4. Accounting for Understanding

The field would benefit from a more holistic approach to understanding that involves the interaction of multiple components or systems. To illustrate what such accounts might look like, we describe conceptual and computational approaches with features of understanding described earlier.

4.1 Conceptual Process Accounts

We first consider two conceptual process models focused on sensemaking (Klein, Moon, & Hoffman, 2006a, 2006b; Klein, Phillips, Rall, & Peluso, 2007; Pirolli & Card, 2005). These accounts describe processes that start at information gathering and end with a degree of understanding about the topic, although they are not implemented as running computational systems. They are not capable of doing things and directly demonstrating their understanding, but they are influential process accounts that deserve attention.

Pirolli and Card's (2005) **sensemaking model** is a dynamic, high level approach that includes a sequence of processes. It involves an information foraging (Pirolli & Card, 1999) and sensemaking loop (Russell, Stefik, Pirolli, & Card, 1993). During foraging, an individual engages in search and filtering of information, then applies an iterative process to give it more structure. Here, there is a tradeoff between exploration and exploitation. During sensemaking, one utilizes schemas to make hypotheses and conclusions, similar to the construction of a mental model (e.g., Johnson-Laird, 1980, 2013), and may recognize a purpose or frame of reference. Problem structuring, reasoning with evidence, and decision making serve as leverage points. This suggests a connection to the higher-order cognitive skills highlighted by educational psychologists.

In contrast, Klein et al.'s (2006a, 2006b, 2007) **data/frame model** is more related to expert analysis and naturalistic decision making. The frame is at the core of sensemaking and determines which information in the environment is deemed relevant, how that information will be interpreted, and what its pragmatic role may be. The sensemaking process is dynamic rather than sequential, however, it typically starts with data and either selection or construction of a frame depending on salient features or anchors in the data. Next, the frame can be further elaborated, questioned, compared with other frames, reframed, discarded, or preserved. Each activity is accomplished through a series of several actions, such as seeking and inferring data, and extending a frame by adding and filling slots (elaboration). The difference between an expert and novice is the quality of a frame, not more advanced reasoning. Often, simple frames are used over more comprehensive ones because they are more generalizable, more efficient, and more pragmatic.

4.2 Computational Process Accounts

In addition to conceptual process accounts, there are several computational-based approaches focused on understanding-related processes and their interaction.

The Construction-Integration Model (Kintsch, 1988) has been applied to text comprehension tasks and differences between novices and experts. During the construction phase, a network of related concepts is formed based on activation of knowledge about the world, as well as semantic and syntactic knowledge. This involves a literal representation of the text and serves as a basic scaffolding. During integration, constraints and context are introduced and influence the activation of the concepts to create an overall representation of the text. Here, general and personal knowledge fill in the scaffolding from the construction phase by adding nodes and relational links. The result of this process is some degree of understanding, which involves incrementally building and organizing knowledge with increasingly abstract relations.

Interactive task learning (Gluck & Laird, 2019-a; Kirk & Laird, 2014; Laird et al., 2017; Mohan, Kirk, & Laird, 2016) and **apprentice learning** (MacLellan, 2017; MacLellan et al., 2016) seek to avoid task specific programming by developing systems that can take instructions, learn basic concepts, and construct more complex concepts. Instruction involves context specific interaction between a learner, which interprets, applies, analyzes, and asks questions, and an instructor, which identifies relevant information and structures tasks. Organizing information into relational networks, using metacognitive monitoring to identify knowledge gaps, and grounded co-construction of the learning experience are all important for interactive task learning. MacLellan et al. (2018) suggest we should focus on training and proposed a natural training interactions framework to guide this process. Agent and instructor interaction patterns (e.g., passive, collaborative, operant conditioning), instructional types (e.g., commands, clarification, demonstration), and various modalities could be used to help develop a natural, common, domain-general method to improve interaction efficiency. These interactive learning approaches relate to understanding. They are dynamic and emphasize relational networks, metacognition, and leveraging context or frames.

The **Companion cognitive architecture** (Forbus & Hinrichs, 2006, 2017; Hinrichs & Forbus, 2014) emphasizes two important components of understanding: multiple interacting components acting as a system and analogical reasoning. Companions has been used to model analogical reasoning in several contexts, such as problem solving in physics (Klenk & Forbus, 2007) Although analogical reasoning is a core component of the architecture, it has other capabilities. For instance, Blass and Forbus (2017) demonstrated some commonsense reasoning capabilities in a companion. The companion started with some commonsense knowledge gained through previous experience. After reading descriptions in natural language, the companion extracts semantic representations and then uses analogical chaining to make inferences. The companion evaluates its inferences and stops processing once an appropriate inference is drawn that aligns with constraints. In addition to multiple interacting components and analogical reasoning, we see a metacognitive process that corresponds to assessing the current degree of understanding and evaluates if it is satisfactory.

DIARC (Schermerhorn et al., 2006; Scheutz, 2006) is a component-based architecture capable of introspection of current states, component sharing across multiple agents, and generating universal representations. These capabilities were leveraged in an introspective three-level framework (Krause, Schermerhorn, & Scheutz, 2012) demonstrating that robots' awareness of physical embodiment and affordances enable them to work together to complete tasks (Scheutz et al., 2019). For instance, in one demonstration, a robot walking backwards was informed to stop by a second observing robot that determined continuing to walk backwards was not safe. This required the second robot to be aware of its own capabilities and constraints and those for the other robot (e.g., can it see?), which involves metacognition and abstract representation.

Laird and Mohan (2018) discussed a dual process-based approach to autonomous agents illustrated with **Rosie**. Implicit architectural learning mechanisms account for System 1 behavior, while deliberation, utilizing knowledge, control, and metacognition accounts for System 2 behavior. Shifting between systems appears to result from metacognitive monitoring. Others call this *reactive* or *reflective* processing and suggest it is an important component in dynamic behavior (Langley et al., 2009; Shanahan, 2006). Recently, Larue, Hough, and Juvina (2018) combined the literature on metacognition, and reactive and reflective processing to implement a metacognitive component into the ACT-R architecture to account for behavior in the Wason card selection task. This model uses a metacognitive component roughly corresponding to one's amount of confidence with a response or outcome, which determines whether a reactive solution is sufficient and if not, reflective processing is triggered to generate an alternative solution. These approaches emphasize the role of metacognition in understanding and provide examples of when and how it occurs.

Thomson and colleagues (Lebiere et al., 2013; Thomson et al., 2012) implemented Pirolli and Card's (2005) sensemaking model into **ACT-R**. Here, they focused on frames, the generation of hypotheses, the search for additional data, and reframing after collecting evidence. The model learned frames or features of objects during a training phase, then during a later phase it was reinforced for utilizing efficient features and sensemaking processes. During a sensemaking task, the model switches between foraging and sensemaking loops to gather evidence, adjust frames, and update hypotheses until an expected utility threshold for making a decision is reached, much like the metacognitive component described in reactive and reflexive processing. Similar to understanding, the model incrementally acquires and organizes knowledge, then builds relations to generate hypotheses. A conclusion is generated and evaluated based on internal or external feedback, which informs whether it is sufficient or further evidence is needed.

The **Metacognitive Integrated Dual-Cycle Architecture** (Cox et al., 2016; Dannenhauer et al., 2018; Paisner, Cox, Maynord, & Perlis, 2014) is a metacognition-focused (i.e., cognition about executed actions and observations regarding objects, object states, and relations), goal-driven system that can solve problems and autonomously regulate its own goals. When discrepancies between metacognitive expectations and the environment are detected, attention is focused towards problem solving rather than the environment and specific actions. Although still under development, it has been applied to some dynamic tasks. For instance, in two different tasks (i.e., Nbeacons and aggressive arsonist), the architecture adapts knowledge or cognitive processing to resolve discrepancies with metacognitive expectations (Dannenhauer et al., 2018). It is also being applied to anticipatory thinking (Amos-Binks & Dannenhauer, 2019), which requires a degree of understanding and recognizing similarity between the past and present in order to predict the future.

5. Toward a Better Understanding of Understanding

In an attempt to synthesize the preceding review, we propose a general definition for the dynamic process and outcome of understanding:

> *The acquisition, organization, and appropriate use of knowledge to produce a response directed towards a goal, when that action is taken with awareness of its perceived purpose.*

During this process, metacognitive monitoring allows identification of knowledge gaps or faulty processing based on detection of discrepancies between knowledge or expectations and the environment. One's current purpose or goal may arise endogenously or exogenously, and in dynamic environments they may change over time. When a clear goal or purpose exists, the depth and breadth of the understanding process can be constrained, and therefore be more efficient.

This definition of understanding clearly places it beyond the capacities of purely reactive controllers and out of reach for any framework, theory, or model that lacks explicit goals. Our conceptualization of understanding relates more closely to Moore and Newell (1974) and Newell (1980), where understanding or rational behavior involves adaptive selection of appropriate actions in order to achieve a goal, reach an end state, or solve a problem. We emphasize that understanding involves metacognitive awareness of purpose and is directed towards goals that may vary in clarity and the degree to which they are satisfied. In addition, we view understanding as being involved in selecting a strategy or approach to solving a problem, and it may inform and be informed by problem solving, but it is not synonymous with problem solving.

The extent of understanding can be conceived of as a network of representations around an initial knowledge base, including how it was learned and its intended use. Understanding increases as knowledge is abstracted, relations are developed, and content is transferred to progressively more distant situations, resulting in higher quality outcomes. The degree of understanding required at a point in time depends on the context, the purpose, and the degree of autonomy. For instance, in some situations, cognitive systems may not need to understand very much to perform effectively, especially when completing simple tasks in low-consequence environments. However, if the system has autonomous control of itself and others in high-risk, dynamically changing situations, then greater understanding can be crucial. Plans or goals are often created prior to task engagement, which makes detecting and recognizing when to switch strategies, modify goals, or engage in deeper processing important.

Measuring understanding is certainly a challenge. The type of information, context, purpose, domain, individual's motivation, and higher-order skills must all be considered. A general and parsimonious method could score features of understanding based on satisfaction, weighting them based on importance for a given context, and then summing them to get a composite score. However, because understanding is complex and there are several measurement options, it is better to measure multiple types of performance, including higher-order cognitive skills, and use higher-density process methods like eye tracking, verbal protocols, and brain imaging. Any attempt at a more complete and holistic assessment of degree of understanding will require an aggregate composite over these kinds of measures and methods.

There are still outstanding issues in understanding-related models with regard to pace, persistence, and partnering (Gluck & Laird, 2019b). On a related note, MacLellan et al. (2018) suggest ways to frame interactions in order to promote learning and understanding. Forbus and Hinrichs (2017) suggest running models for longer periods of time to allow construction of knowledge structures that can later answer questions or solve problems, thus aiding scaling and longevity. We have provided a foundation on which to build cognitive systems that understand and to measure the extent to which they achieve this aim. It is also clear that challenges remain in our ability to create systems with this capacity. We hope this review promotes additional constructive progress.

Acknowledgements

This research was supported in part by Alex Hough's appointment to the Student Research Participation Program at the U.S. Air Force Research Laboratory, 711th Human Performance Wing, Cognitive Science, Models, and Agents Branch, administered by the Oak Ridge Institute for Science and Education. We thank the Chief Scientist's office at the 711th Human Performance Wing for support through the Repperger summer internship program. We thank Leslie Blaha, two anonymous reviewers, and Pat Langley for constructive feedback and suggestions on earlier drafts.

References

Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamin, R., & Zhang, D. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of Educational Research, 78*, 1102–1134.

Aha, D. W. (2018). Goal reasoning: Foundations, emerging applications, and prospects. *AI Magazine,* 39, 3–24.

Anderson, J. A., Reder, L. M., & Simon, H. A. (1997). Situative versus cognitive perspectives: Form versus substance. *Educational Researcher*, 26, 18–21.

Anderson, J. R., & Schunn, C. D. (2000). Implications of the ACT-R learning theory: no magic bullets. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 5), 1–33). Mahwah, NJ: Lawrence Erlbaum.

Anderson, J. R., Greeno, J. G., Reder, L. M., & Simon, H. A. (2000). Perspectives on learning, thinking, and activity. *Educational Researcher*, *29*, 11–13.

Austin, J. L. (1962). *How to do things with words*. Cambridge, MA: Harvard University Press.

Ball, J. (2008). A naturalistic, functional approach to modeling language comprehension. *Papers from the AAAI Fall 2008 Symposium on Naturally Inspired Artificial Intelligence* (pp.1–8). Arlington, VA: AAAI Press.

Ball, J., Myers, C., Heiberg, A., Cooke, N. J., Matessa, M., Freiman, M., & Rogers, S. (2010). The synthetic teammate project. *Computational and Mathematical Organization Theory*, 16, 271–299.

Berardi-Coletta, B., Buyer, L. S., Dominowski, R. L., & Rellinger, E. R. (1995). Metacognition and problem solving: A process-oriented approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 205–223.

Blass, J. A., & Forbus, K. D. (2017). Analogical chaining with natural language instruction for commonsense reasoning. *Proceedings of the Thirty-first AAAI Conference on Artificial Intelligence* (pp. 4357–4363). San Francisco, CA: AAAI Press.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11*, 717–726.

Brenner, M. E., Mayer, R. E., Moseley, B., Brar, T., Durán, R., Reed, B. S., & Webb, D. (1997). Learning by understanding: The role of multiple representations in learning algebra. *American Educational Research Journal*, *34*, 663–689.

Bunge, S. A. (2004). How we use rules to select actions: A review of evidence from cognitive neuroscience. *Cognitive, Affect, and Behavioral Neuroscience*, 4, 564–579.

Bunge, S. A., Wendelken, C., Badre, D., & Wagner, A. D. (2005). Analogical reasoning and prefrontal cortex: Evidence for separable retrieval and integration mechanisms. *Cerebral Cortex, 15*, 239–249.

Butcher, K., & Sumner, T. (2011). Self-directed learning and the sensemaking paradox. *Human-Computer Interaction*, 26, 123–159.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55–81.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. S. Sternberg (Ed.). *Advances in the psychology of human intelligence* (Vol. 1), 1–75. Hillsdale, NJ: Lawrence Erlbaum

Clarke, D., Cheeseman, J., Roche, A., & van der Schans, S. (2014). Teaching strategies for building student persistence on challenging tasks: Insights emerging from two approaches to teacher professional learning. *Mathematics Teacher Education and Development*, 16, 46–70.

Cox, M. T. (2007). Perpetual self-aware cognitive agents. *AI Magazine*, 28, 32–45.

Cox, M. T., Alavi, Z., Dannenhauer, D., Eyorokon, V., Muñoz-Avila, H., & Perlis, D. (2016). MIDCA: A metacognitive, integrated dual-cycle architecture for self-regulated autonomy. *Proceedings of the Thirteenth AAAI Conference on Artificial Intelligence* (pp. 3712–3718). Phoenix, AZ: AAAI Press.

Cunningham, J. B., & MacGregor, J. N. (2014). Productive and re-productive thinking in solving insight problems. *The Journal of Creative Behavior*, *48*, 44–63.

Dannenhauer, D., Cox, M. T., & Muñoz-Avila, H. (2018). Declarative metacognitive expectations for high-level cognition. *Advances in Cognitive Systems*, 6, 231–250.

de Groot, A. D. (1965). *Thought and choice in chess*. Cambridge, UK: Mouton Publishers.

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, 106, 1248–1299.

Duncker, K. (1945). On problem solving. *Psychological Monographs*, 58, 1–113.

Eliasmith, C., & Thagard, P. (2001) Integrating structure and meaning: A distributed model of analogical mapping. *Cognitive Science, 25*, 245–286

Ennis, R. H., & Millman, J. (1985). *Cornell Critical Thinking Test, Level X*. Pacific Grove, CA: Midwest.

Ericsson, A., Polson, P. G., & Wertheimer, M. (1982). Preface to the Phoenix edition. In M. Wertheimer (Ed.), *Productive thinking*, 11–17. Chicago, IL: University of Chicago Press.

Ericsson, K. A., & Harris, M. S. (1990). *Expert chess memory without chess knowledge: A training study*. Poster presented at the Thirty-Ninth Annual Meeting of the Psychonomic Society. New Orleans, LA.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, *47*, 273–305.

Ericsson, K. A., & Oliver, W. L. (1989). A methodology for assessing the detailed structure of memory skills. In A. M. Colley & J. R. Beech (Eds.), *Acquisition and performance of cognitive skills*, 193–215. Chichester, UK: Wiley.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*, 223–241.

Facione, P. A. (1990). *The California Critical Thinking Skills Test – College Level: Experimental validation and content validity* (Technical Report No. 1). Millbrae, CA: Calif Academic Press.

Forbus, K. D., & Hinrichs, T. R. (2006). Companion cognitive systems: A step toward human-level AI. *AI Magazine, 27*, 83–95.

Forbus, K. D., & Hinrichs, T. R. (2017). Analogy and qualitative representations in the Companion cognitive architecture. *AI Magazine, 38*, 34–42.

Forbus, K., & Gentner, D. (1997). Qualitative mental models: Simulations or memories. *Proceedings of the Eleventh International Workshop on Qualitative Reasoning* (pp. 3–6). Cortona, Italy.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*, 25–42.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*, 155–170.

Gettier, E. L. (1963). Is justified true belief knowledge? *Analysis*, 23, 121–123.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*, 1–38.

Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory & Cognition, 10*, 597–602.

Gluck, K. A., & Laird, J. E. (Eds.) (2019a). *Interactive task learning*. Cambridge, MA: MIT Press.

Gluck, K. A., & Laird, J. E. (2019b). Looking forward to interactive task learning. In K. A. Gluck & J. E. Laird (Eds.), *Interactive task learning*, 1–7. Cambridge, MA: MIT Press.

Gluck, K. A., Staszewski, J. J., Richman, H., Simon, H. A., & Delahanty, P. (2001). The right tool for the job: Information-processing analysis in categorization. *Proceedings of the Twenty-Third Annual Meeting of the Cognitive Science Society* (pp. 330–335). London: Lawrence Erlbaum.

Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, *31*, 1–40.

Green, A. E., Fugelsang, J. A., Kraemer, D. J., Shamosh, N. A., & Dunbar, K. N. (2006). Frontopolar cortex mediates abstract integration in analogy. *Brain Research*, 1096, 125–137.

Greeno, J. G. (1977). Process of understanding in problem solving. In N. J. Castellan, D. B. Pisoni, & G. R. Potts (Eds.), *Cognitive theory* (Vol 2). Hillsdale, NJ: Lawrence Erlbaum.

Greeno, J. G. (1997). On claims that answer the wrong questions. *Educational Researcher*, *26*, 5–17.

Greeno, J. G., Riley, M. S., & Gelman, R. (1984). Conceptual competence and children's counting. *Cognitive Psychology, 16*, 94–143.

Grimm, S. R. (2006). Is understanding a species of knowledge? *British Society for the Philosophy of Science, 57*, 515–535.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American Psychologist*, *53*, 449–455.

Halpern, D. F. (2010). Halpern Critical Thinking Assessment. Mödling, Austria: SCHUHFRIED. http://www.schuhfried.com/vienna-test-system-vts/all-tests-from-a-z/test/heta-halpern-critical-thinking-assessment-1/

Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking*. New York, NY: Psychology Press.

Hayes, J. R., & Simon, H. A. (1974). Understanding written problem instructions. In L. W. Gregg (Ed.), *Knowledge and cognition*. Potomac, MD: Lawrence Erlbaum.

Hinrichs, T. R., & Forbus, K. D. (2014). X goes first: Teaching simple games through multimodal interaction. *Advances in Cognitive Systems, 3*, 31–46.

Hofstadter, D. R. (2001). Analogy as the core of cognition. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.) *The analogical mind: Perspectives from cognitive science*, 499–538. Cambridge, MA: MIT Press.

Hummel, J. E., & Holyoak, K. J. (2005). Relational reasoning in a neurally plausible cognitive architecture: An overview of the LISA project. *Current Directions in Psychological Science, 14*, 153–157.

Ischebeck, A., Zamarian, L., Schocke, M., & Delazer, M. (2009). Flexible transfer of knowledge in mental arithmetic—An fMRI study. *Neuroimage*, *44*, 1103–1112.

Johnson, B., Roberts, M., Apker, T., & Aha, D. W. (2016). Goal reasoning with information measures. *Advances in Cognitive Systems*, 4, 1–14.

Johnson-Laird, P. N. (2013). The mental models perspective. In D. Reisberg (Ed.), *The Oxford handbook of cognitive psychology*, 650–667. New York: Oxford University Press.

Kant, I. (1958). *Critique of pure reason* (trans. by N. Kemp Smith). New York: Random House. Original work published in 1781.

Katona, G. (1940). *Organizing and memorizing*. New York: Columbia University Press.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95, 163–182.

Kirk, J. R., & Laird, J. (2014). Interactive task learning for simple games. *Advances in Cognitive Systems, 3*, 13–30.

Klein, G., Moon., B., & Hoffman, R. R. (2006a). Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems*, 21, 70–73.

Klein, G., Moon., B., & Hoffman, R. R. (2006b). Making sense of sensemaking 2: A macro cognitive model. *IEEE Intelligent Systems*, 21, 88–92.

Klein, G., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2007). A data-frame theory of sensemaking. *Proceedings of the Sixth International Conference on Naturalistic Decision Making* (pp. 113–155). New York, NY: Lawrence Erlbaum.

Klenk, M., & Forbus, K. (2007). Cognitive modeling of analogy events in physics problem solving from examples. *Proceedings of the Twenty-ninth Annual Conference of the Cognitive Science Society* (pp. 1153–1158). Nashville, TN.

Klenk, M., Molineaux, M., & Aha, D. W. (2013). Goal-driven autonomy for responding to unexpected events in strategy simulations. *Computational Intelligence*, *29*, 187–206.

Kohler, W. (1929). *Gestalt psychology*. New York, NY: Liveright.

Krause, E. A., Schermerhorn, P., & Scheutz, M. (2012). Crossing boundaries: Multi-level introspection in a complex robotic architecture for automatic performance improvements. *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence* (pp. 214–220). Toronto, ON: AAAI Press.

Krawczyk, D. C. (2010). The cognition and neuroscience of reasoning. *Brain Research, 1428*, 14–23.

Laird, J. E., & Mohan, S. (2018). Learning fast and slow: Levels of learning in general autonomous intelligent agents. *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence* (pp. 7983–7987). New Orleans, LA: AAAI Press.

Laird, J. E., Gluck, K., Anderson, J., Forbus, K. D., et al. (2017) Interactive task learning. *IEEE Intelligent Systems, 32*, 6–21.

Langley, P., & Jones, R. (1988). A computational model of scientific insight. In R. J. Sternberg (Ed.), *The nature of creativity*, 177–201. New York, NY: Cambridge University Press.

Langley, P., Laird, J. E., & Rogers, S. (2009). Cognitive architectures: Research issues and challenges. *Cognitive Systems Research*, *10*, 141–160.

Larkin, J. H., McDermott, J. Simon, D. P., Simon, H. A. (1980). Models of competence in solving physics problems. *Cognitive Science, 4*, 317–345.

Larue, O., Hough, A., & Juvina, I. (2018). A cognitive model of switching between reflective and reactive decision making in the Wason task. *Proceedings of the Sixteenth International Conference on Cognitive Modeling* (pp. 55–60). Madison, WI: University of Wisconsin.

Lebiere, C., Pirolli, P., Thomson, R., Paik, J., Rutledge-Taylor, M., Staszewski, J., & Anderson, J. R. (2013). A functional model of sensemaking in a neurocognitive architecture. *Computational Intelligence and Neuroscience, 5*, 1–29.

Lindsay, R. K. (1963) Inferential memory as the basis of machines which understand natural language. In E. A. Feigenbaum & J. Feldman, *Computers and thought,* 217–236. New York, NY: McGraw-Hill.

MacLellan, C. J. (2017). *Computational models of human learning: Applications for tutor development, behavior prediction, and theory testing*. Doctoral dissertation, Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA.

MacLellan, C. J., Harpstead, E., Marinier III, R. P., & Koedinger, K. R. (2018). A framework for natural cognitive system training interactions. *Advances in Cognitive Systems*, *6*, 177–192.

MacLellan, C. J., Harpstead, E., Patel, R., & Koedinger, K. R. (2016). The Apprentice Learner architecture: Closing the loop between learning theory and educational data. *Proceedings of the Ninth International Conference on Educational Data Mining* (pp. 151–158). Boston, MA: International Education Data Mining Society.

Maier, N. R. F. (1931). Reasoning in humans: II. The solution of a problem and its appearance in consciousness. *Journal of Comparative Psychology, 12,* 181–194.

Mannes, S. M., & Kintsch, W. (1987). Knowledge organization and text organization. *Cognition and Instruction, 4*, 91–115.

Marshall, J. B., & Hofstader, D. R. (1997). The Metacat project: A self-watching model of analogy making. *Cognitive Studies*, 4, 57–71.

Mata, A., Ferreira, M. B., & Sherman, S. J. (2013). The metacognitive advantage of deliberative thinkers: A dual-process perspective on overconfident. *Journal of Personality and Social Psychology*, *105*, 353–373.

Mata, A., Ferreira, M. B., Voss, A., & Kollei, T. (2017). Seeing the conflict: An attentional account of reasoning errors. *Psychonomic Bulletin & Review*, 24, 1980–1986.

Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional Science*, *26*, 49–63.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*, 1–43.

Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision*, 211–277. New York, NY: McGraw-Hill.

Mohan, S., Kirk, J., & Laird, J. (2013). A computational model for situated task learning with interactive instruction. *Proceedings of the Thirteenth International Conference on Cognitive Modeling*. Ottawa, Canada.

Moore, J., & Newell, A. (1974). How can Merlin understand? In L. W. Gregg (Ed.), *Knowledge and cognition*. Potomac, MD: Lawrence Erlbaum.

Muñoz-Avila, H., & Cox, M. T. (2008). Case-based plan adaptation: An analysis and review. *IEEE Intelligent Systems*, 23, 75–81.

Newell, A. (1980). Physical symbol systems. *Cognitive Science*, *4*, 135–183.

Newell, A. (1994). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Newell, A., & Simon, H. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Paisner, M., Cox., M. T., Maynord, M., & Perlis, D. (2014). Goal-driven autonomy for cognitive systems. *Proceedings of the Thirty-Sixth Annual Conference of the Cognitive Science Society* (pp. 2085–2090). Quebec City, Quebec: Cognitive Science Society.

Pennycook, G. (2017). A perspective on the theoretical foundation of dual process models. In W. De Neys (Ed.), *Dual Process Theory 2.0*, 6–26. New York, NY: Routledge.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology*, 80, 34–72.

Perkins, D. N. (1998). What is understanding? In M. S. Wiske (Ed.), *Teaching for understanding: Linking research with practice*, 39–57. San Francisco, CA: Jossey-Bass.

Perkins, D. N., & Simmons, R. (1988). Patterns of misunderstanding: An integrative model for science, math, and programming. *Review of Educational Research*, 58, 303–326.

Pirolli, P., & Card, S. K. (1999). Information foraging. *Psychological Review, 106*, 643–675.

Pirolli, P., & Card, S. K. (2005). The sensemaking process and leverage points for analyst technology. Paper presented at the *International Conference on Intelligence Analysis*, McLean, VA.

Pols, A. J. K. (2002). *Insight problem solving*. Doctoral dissertation, University of Utrecht, Utrecht, Netherlands.

Prat, C. S., Mason, R. A., & Just, M. A. (2012). An fMRI investigation of analogical mapping in metaphor comprehension: The influence of context and individual cognitive capacities on processing demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 282–294.

Pritchard, D. (2009). Knowledge, understanding and epistemic value. *Royal Institute of Philosophy Supplement*, *64*, 19–43.

Pritchard, D. (2016). Seeing it for oneself: Perceptual knowledge, understanding, and intellectual autonomy. *Episteme, 13*, 29–42.

Ram, A., & Leake, D. (1995). *Goal-driven learning*. Cambridge, MA: MIT Press.

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology, 94*, 249–273.

Roberts, M., Borrajo, D., Cox, M. T., & Yorke-Smith, N. (2018). Special issue on goal reasoning. *AI Communications*, 31, 115–116.

Rorty, R. (1979). *Philosophy and the mirror of nature*. Princeton, NJ: Princeton University Press.

Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993) The cost of structure of sensemaking. Paper presented at the *INTERCHI 1993 Conference on Human Factors in Computing Systems*. Amsterdam.

Schank, R. C. (1980). Language and memory. *Cognitive Science, 4*, 243–284.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Lawrence Erlbaum.

Schermerhorn, P., Kramer, J., Brick, T., Anderson, D., Dingler, A., & Scheutz, M. (2006). DIARC: A testbed for natural human-robot interactions. *AAAI Mobile Robot Competition and Exhibition: Papers from the 2006 AAAI Workshop* (Technical Report WS-06-15). Menlo Park, CA: AAAI Press.

Scheutz, M. (2006). ADE: Steps toward a distributed development and runtime environment for complex robotic agent architectures. *Applied Artificial Intelligence, 20*, 275–304.

Scheutz, M., Williams, T., Krause, E., Oosterveld, B., Sarathy, V, & Frasca, T. (2019). An overview of the distributed integrated cognition affect and reflection DIARC architecture. In M. I. A. Ferreira, J. S. Sequeira, & R. Ventura (Eds.), *Cognitive architectures*, 165–193. Cham, Switzerland: Springer.

Schierwagen, A. (2009). Brain complexity: Analysis, models and limits of understanding. In J. Mira, J. M. Ferrandez, J. R. Alverez, F. de la Paz, & F. J. Toledo (Eds.), *Methods and models in artificial and natural computation,* 195–204. Berlin: Springer,

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. Grouws (Ed.) *Handbook for research on mathematics teaching and learning*, 334–370. New York, NY: Macmillian.

Scholz, O. R. (2018). Symptoms of expertise: Knowledge, understanding and other cognitive goods. *Topoi*, *37*, 29–37.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3, 417–457.

Shanahan, M. (2006). A cognitive architecture that combines internal simulation with global workspace. *Consciousness and Cognition*, 15, 433–449.

Shapiro, S., Sardina, S., Thangarajah, J., Cavedon, L., & Padgham, L. (2012). Revising conflicting intention sets in BDI agents. *Proceedings of the Eleventh International Conference on Autonomous Agents and Multiagent Systems* (pp. 1081–1088). Valencia, Spain.

Simms, N. K., Frausel, R. R., & Richland, L. E. (2018). Working memory predicts children's analogical reasoning. *Journal of Experimental Child Psychology*, *166*, 160–177.

Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R. S. Siegler (Ed.), *Children's thinking: What develops?*, 325–348. Hillsdale, NJ: Lawrence Erlbaum.

Simon, H. A. (1980). Information-processing explanations of understanding. In P. W. Jusczyk & R. M. Klein (Eds.), *The nature of thought: Essays in honor of D.O. Hebb*, 37–49. Hillsdale, NJ: Lawrence Erlbaum

Simon, H. A. (1986). The information processing explanation of gestalt phenomena. *Computers in Human Behavior, 2*, 241–255.

Simon, H. A. (1996). *The sciences of the artificial*. Cambridge, MA: MIT Press.

Simon, H. A., & Gilmartin, K. (1973). A simulation of memory for chess positions. *Cognitive Psychology*, *5*, 29–46.

Simon, H. A., & Hayes, J. R. (1976). The understanding process: Problem isomorphs. *Cognitive Psycholog*y, 8, 165–190.

Skyrms, B. (1980). *Casual necessity*. New Haven. CT: Yale University Press.

Smith, M. U., & Siegel, H. (2004). Knowing, believing, and understanding: What goals for science education? *Science & Education, 13*, 553–582.

Stanovich, K. E. (2018). Miserliness in human cognition: The interaction of detection, override and mindware. *Thinking & Reasoning*, 24, 423–444.

Sweller, J., Jeroen, van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251–296.

Thompson, V. A., Evans, J. S. B., & Campbell, J. I. (2013). Matching bias on the selection task: It's fast and feels good. *Thinking & Reasoning, 19*, 431–452.

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, 63, 107–140.

Thomson, R., Lebiere, C., Rutledge-Taylor, M., Staszewski, J., & Anderson, J. R. (2012). Understanding sensemaking using functional architectures. *Proceedings of the Twenty-First Conference on Behavior Representation in Modeling and Simulation.* Amelia Island, FL.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, 39, 1275–1289.

Vattam, S., Klenk, M., Molineaux, M., & Aha, D. W. (2013). Breadth of approaches to goal reasoning: A research survey. *Goal Reasoning: Papers from the ACS Workshop (Technical Report CS-TR-5029)* (pp. 111–126). College Park, MD: University of Maryland.

Veloso, M. M., & Carbonell, J. G. (1993). Derivational analogy in prodigy: automating case acquisition, storage, and utilization. *Machine Learning, 10*, 249–278.

Walsh, M. M., Collins, M., & Gluck, K. A. (2015). Verbal reports reveal strategies in multiple-cue probabilistic inference. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the Thirty-Seventh Annual Conference of the Cognitive Science Society* (pp. 2589–2594). Austin, TX: Cognitive Science Society.

Walsh, M. W., & Gluck, K. A. (2016). Verbalization of decision strategies in multiple-cue probabilistic inference. *Journal of Behavioral Decision Making*, *29*, 78–91.

Watson, G., & Glaser, E. M. (1980). *Watson-Glaser Critical Thinking Appraisal*. San Antonio, TX: PsychCorp.

Wertheimer, M. (1985). A gestalt perspective on computer simulations of cognitive processes. *Computers in Human Behavior*, 1, 19–33.

Whitaker, K. J., Vendetti, M. S., Wendelken, C., & Bunge, S. A. (2018). Neuroscientific insights into the development of analogical reasoning. *Developmental Science*, *21*, e12531.

Wilson, M., Molineaux, M., & Aha, D. W. (2013). Domain-independent heuristics for goal formulation. *Proceedings of the Twenty-Sixth International FLAIRS Conference* (pp. 160–165). St. Pete Beach, FL.

Wittgenstein, L. (1961). *Tractatus logico-philosophicus* (Trans. by D. F. Pears & B. F. McGuinness). London: Routledge & Kegan Paul. Original work published in 1921.

Woodward, J. (2003) *Making things happen: A theory of causal explanation.* New York, NY: Oxford University Press.

Young, J., & Hawes, N. (2012). Evolutionary learning of goal priorities in a real-time strategy game. *Proceedings of the Eighth Annual International Conference on Artificial Intelligence and Interactive Digital Entertainment* (pp. 87–92). Palo Alto, CA.

Zagzebski, L. (2001). Recovering understanding. In M. Steup (Ed.), *Knowledge, truth, and duty: Essays on epistemic justification, responsibility, and virtue*. New York: Oxford University Press.