Measuring Progress Toward AGI: A Cognitive Framework

Ryan Burnell[1], **Yumeya Yamamori**[1], **Orhan Firat**[1], **Kate Olszewska**[1], **Steph Hughes-Fitt**[1], **Oran Kelly**[1], **Isaac R. Galatzer-Levy**[1], **Meredith Ringel Morris**[1], **Allan Dafoe**[1], **Alison M. Snyder**[1], **Noah D. Goodman**[1*], **Matthew Botvinick**[1*] **and Shane Legg**[1]
[1]Google DeepMind, [*]Work done while at Google DeepMind.

Despite widespread discussion of AGI, there is no clear framework for measuring progress toward it. This ambiguity fuels subjective claims, makes it difficult to track progress, and risks hindering responsible governance. As a starting point to address this gap, we present a framework for understanding system capabilities in relation to human cognitive abilities. Drawing from decades of research in psychology, neuroscience, and cognitive science, we introduce a Cognitive Taxonomy that deconstructs general intelligence into 10 key cognitive faculties. We then propose a rigorous evaluation protocol in which a system's performance is measured across a suite of targeted, held-out cognitive tasks, generating a 'cognitive profile' that can be used to understand a system's strengths and weaknesses. We hope this framework will provide a practical roadmap and an initial step toward more rigorous, empirical evaluation of AGI.

1. Introduction

Artificial general intelligence (AGI) has the potential to accelerate scientific discovery, increase productivity and help solve some of humanity's most pressing problems. Yet our ability to understand how close we are to this critical milestone is hampered by a lack of clarity about how general intelligence should be operationalized and measured. As a result, the capabilities of today's AI tools are often underestimated or overhyped while the potential benefits and risks of future systems remain poorly understood. This lack of empirical grounding makes it difficult for researchers to communicate progress and for policymakers to craft effective governance, ultimately hindering our collective ability to navigate the path to AGI responsibly. To address this critical gap, we introduce a cognitive framework for measuring progress on increasingly capable AI systems.

1.1. Operationalizing AGI

The term AGI has a long history. It was first used by Mark Gubrud in a 1997 paper to refer to systems that "*rival or surpass the human brain in complexity and speed, that can acquire, manipulate and reason with general knowledge, and that are usable in essentially any phase of operations where a human intelligence would otherwise be needed*" (Gubrud, 1997). Then, in 2001, without knowledge of Gubrud's work, Shane Legg independently coined the term "Artificial General Intelligence" for Ben Goertzel's subsequent book, which adopted it as its title (Goertzel and Pennachin, 2007). There, AGI was referred to as "*AI systems that possess a reasonable degree of self-understanding and autonomous self-control, and have the ability to solve a variety of complex problems in a variety of contexts, and to learn to solve new problems that they didn't know about at the time of their creation*". Since then, AGI has become a mainstay in discussions about AI capabilities. However, the term is often used as a shorthand to describe various kinds of highly capable AI systems. Given the important societal and scientific implications of generally capable AI systems, it is vital that we establish precise and robust ways to measure progress toward this milestone.

Corresponding author(s): rburnell@google.com
© 2026 Google. All rights reserved

As a first step in this direction, in 2023 Google DeepMind presented the *Levels of AGI* framework that proposed a series of important stages on the path toward general intelligence (Morris et al., 2024). The framework considers intelligence a continuous, multidimensional construct and argues it is important for systems to be both highly *capable* and highly *general*. On both these dimensions, it is clear that human capabilities are a key reference point. The creation of a system that is capable of exhibiting all the cognitive capabilities that humans have would mark a historic moment and a philosophical milestone. Moreover, such a system would open up countless applications, including universal assistants, personalized learning, and powerful new scientific tools.

What remains less clear is how we can understand how far away AI systems are from matching these cognitive capabilities. Over the past few years, several benchmarks aimed at measuring progress toward AGI have been proposed (see, for example, Chollet 2019 and Hendrycks et al. 2025). However, existing efforts fail to cover the full breadth of human cognition and lack robust comparisons to human performance. Here, we aim to address this gap in two parts. First, we propose a Cognitive Taxonomy that captures the important aspects of human cognition that an AGI system should be able to match. Second, we describe a framework for evaluating AI systems across this cognitive space to help us better contextualize system capabilities.

2. Cognitive Taxonomy

To understand where AI systems stand relative to human cognitive capabilities, we first need to identify the key cognitive processes that enable people to navigate the complex and changing world. In this endeavor, we can turn to the decades of research into human cognition—in particular, research from *psychology, neuroscience, and cognitive science*, which have built and refined theories of cognition over many decades through iterative experimentation. These disciplines employ a wide range of methods, including experimental paradigms, brain imaging techniques, patient studies, and computational modeling, providing rich and empirically grounded insights we can draw from. Here, we use these insights to create a *cognitive taxonomy* describing the cognitive capabilities that evidence suggests are important for general intelligence.

Of course, characterizing intelligence is far from a simple task. Even within the field of human cognitive science there are many debates that remain unsettled and many questions that remain unanswered (Stainton, 2006). The world of artificial systems is perhaps even more complex, rich with a multitude of architectures and training algorithms improving at a pace that makes evolution pale in comparison. It is therefore possible—even likely—that some aspects of human cognition may not be relevant in the context of artificial systems. Conversely, we are likely to find cognitive processes in these artificial systems that humans do not possess.

For this reason, our goal was not to create *the* definitive account of all of cognition, but rather to build a practical framework for evaluation that is both theoretically grounded and comprehensive enough to cover the breadth of human intelligence. We consider this framework a starting point, and we hope it can provide a foundation for a robust science of artificial general intelligence.

2.1. Cognitive faculties

Our cognitive taxonomy enumerates 10 cognitive faculties that the scientific efforts in mapping human cognition suggest are important for intelligent behavior (see Figure 1). For each faculty and process, we identify a set of specific abilities and sub-abilities that a generally intelligent system should be expected to exhibit. For brevity, we provide brief descriptions of each faculty below and include the full taxonomy in the Appendix.

An important characteristic of our taxonomy is that it focuses on *what* the system is able to accomplish not on *how* it does so (see Marr 1982 for a discussion of this distinction). In doing so, we are able to remain agnostic to the underlying mechanisms employed by a system and do not prescribe any specific modeling approaches.

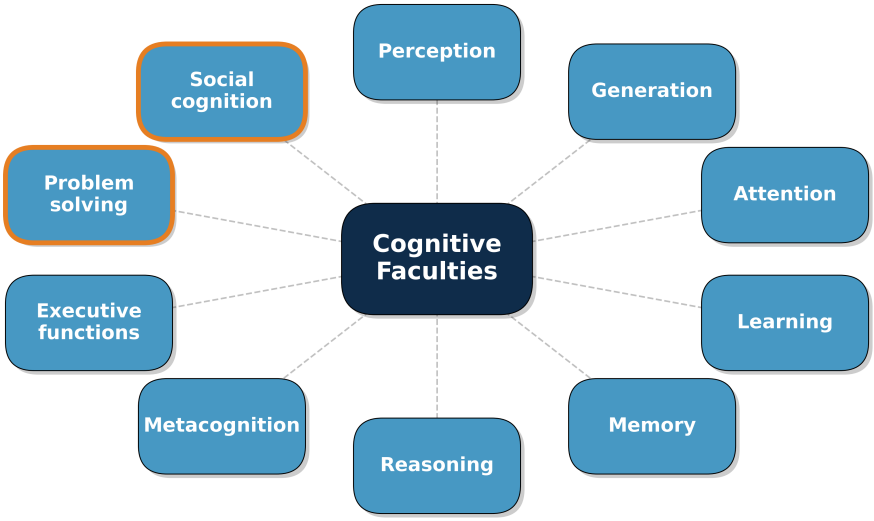


Figure 1 | *Overview of the 10 cognitive faculties. Faculties outlined in orange represent composite faculties.*

We begin with eight faculties capturing the basic building blocks of human cognition:

Perception: The ability to extract and process sensory information from the environment.

Generation: The ability to produce outputs such as speech, text, motor movements, and computer control actions.

Attention: The ability to focus cognitive resources on specific aspects of perceptual stimuli, thoughts, or task demands.

Learning: The ability to acquire new knowledge, skills, or understanding through experience, study, or instruction.

Memory: The ability to store and retrieve information over time.

Reasoning: The ability to draw valid conclusions and make inferences by applying logical principles.

Metacognition: The knowledge a system has about its own cognitive processes and its ability to monitor and control those processes.

Executive functions: Abilities that facilitate goal-directed behavior. Includes planning, inhibition, and cognitive flexibility.

These eight faculties form the basic building blocks, but they do not operate in isolation—on the contrary, they interact, overlap, work together, and build on one another (Kovacs and Conway, 2016). To fully understand the capabilities of a system, we also need to understand how well it can combine

and apply multiple faculties in unison. We therefore propose an additional two composite faculties capturing two critical psychological contexts in which the various faculties are applied together:

Problem solving: The ability to find effective solutions to domain-specific problems.

Social cognition: The ability to process and interpret social information and to respond appropriately in social situations.

We hypothesize that these ten faculties capture the key capabilities needed for a system to be both highly general and highly capable. A system with significant weaknesses in one or more of these ten faculties is likely to be unable to perform some real-world tasks that most humans can perform.

3. Evaluating Cognitive Capabilities

To understand the cognitive capabilities of an AI system, we need to robustly evaluate the system across each cognitive faculty and compare the system to a meaningful human baseline. We therefore propose the following three-stage evaluation protocol:

1. **Conduct cognitive assessment** of the system across a broad suite of cognitive tasks.
2. **Collect human baselines** on the same cognitive tasks to establish a point of comparison.
3. **Build cognitive profiles** to map the system's strengths and weaknesses in relation to human performance.

3.1. Conduct cognitive assessment

The first step in evaluating progress toward AGI is to **evaluate system performance on a broad suite of cognitive tasks** covering each faculty. The cognitive tasks should be:

- **Targeted to specific cognitive abilities**. It is important to include tasks that isolate each cognitive faculty in order to precisely diagnose model weaknesses.
- **Held out**. Evaluations should ideally use private, held-out test sets to prevent contamination—if systems had previously seen solutions or strategies for the specific test items, then performance on those tests is unlikely to be indicative of general intelligence (e.g., see Jacovi et al. 2023).
- **Independently verified**. To ensure the community can be confident in the findings, both the cognitive tasks and the evaluation results should be audited by an independent third party.
- **Varied in difficulty for humans**. Some tasks are easy for humans and hard for AI systems (see Chollet et al. 2025), so it is important to include such tasks as well as tasks that test the limits of human capabilities.
- **Varied in structure and format**. Idiosyncrasies in the structure of specific tasks can artificially inflate (or inhibit) performance on those tasks (this is true both for humans and AI systems; e.g. see Cheng and Holyoak 1985; Dasgupta et al. 2024). For this reason, the evaluation of each faculty should include multiple tasks with a variety of structures and formats (e.g. multiple choice vs open response, text inputs vs multimodal, multi-step vs single-turn).

3.2. Collect human baselines

To understand how close a system's capabilities are to human-level on a set of tasks, we need to **quantify human performance** on those same tasks.

These *human baselines* can be constructed by asking a large sample of humans to complete the same tasks as the AI systems. The tasks should be performed under the same conditions, including the same task instructions (and few-shot examples, if any), response format, and access to external tools.

Because we want to understand the full range of human capabilities, we think it is critical to sample widely from the human population. At the same time, we want to understand how well systems will perform in real-world situations—situations that involve knowledge and abilities that are typically only fully developed in adulthood and typically honed through formal education. Therefore, we propose that a reasonable human baseline should consist of a **demographically representative** sample of **adults** with at least the equivalent of an **upper secondary education**[1].

3.3. Build cognitive profiles

Using the evaluation results from the cognitive assessment and the human baselines, we can next **build cognitive profiles** for each system to map out the strengths and weaknesses of the system relative to human performance across the 10 cognitive faculties.

To do so, we can calculate the percentage of people from the human sample the system outperforms, and place the system along the distribution of human performance (see Figure 2 for examples).

Given the jagged nature of system capabilities (Morris et al., 2026), many kinds of profiles are possible. For example:

1. A system scores **below the human baseline sample median** for **one or more** cognitive faculties. Such a system shows significant cognitive weaknesses and is likely to struggle in at least some real-world contexts.
2. A system scores **above the human baseline sample median** across **all ten** cognitive faculties. Such a system is demonstrating the ability to match the cognitive capabilities of at least 50% of the human sample, and is likely to perform well in many real-world contexts.
3. A system scores in the **99th percentile** across **all ten** cognitive faculties (Figure 2C). Such a system is demonstrating it can match almost anyone from the human sample across the entire cognitive space.[2] Of course, any practical human baseline sample—even a highly representative one—is unlikely to capture the full scope of human capabilities, so further scrutiny would be required before drawing any conclusions based on this pattern.

Some important methodological notes are warranted here: To construct a cognitive profile, we need to calculate an overall score for each person and each AI system on each faculty (see Figure 2). There are many ways these scores could be calculated. The simplest approach would be to aggregate together the metrics from each cognitive task within a faculty. But taking a more sophisticated statistical approach—such as building an Item-Response Theory model of system capabilities (see e.g., Martínez-Plumed et al. 2019)—could provide more robust and informative results.

Regardless of which approach is used, it is important to quantify the uncertainty around each capability estimate. Uncertainty in this context stems from at least three main sources:

1. **Task quality**: Factors such as prompt quality and task diversity can dramatically affect how informative the results are.

[1] Equivalent to earning a high-school degree in the US.

[2] A somewhat stronger test would be to require the system to score at or above the sample maximum across every faculty. But it is difficult to quantify the uncertainty around a maximum value.

2. **Construct validity**: Cognitive evaluations are intended to provide insights into broad cognitive faculties, such as reasoning or attention. However, if the datasets do not isolate the target faculty effectively or the data are contaminated, the results may not accurately reflect a system's general capabilities.
3. **Stochasticity**: The stochasticity of generative AI systems adds noise to evaluation results—asking a system to complete the same task multiple times can produce wildly different results across repetitions (this is particularly true of complex tasks with multiple steps).

Characterizing this uncertainty is crucial for understanding the extent to which differences between systems or between a system and the human sample are meaningful.

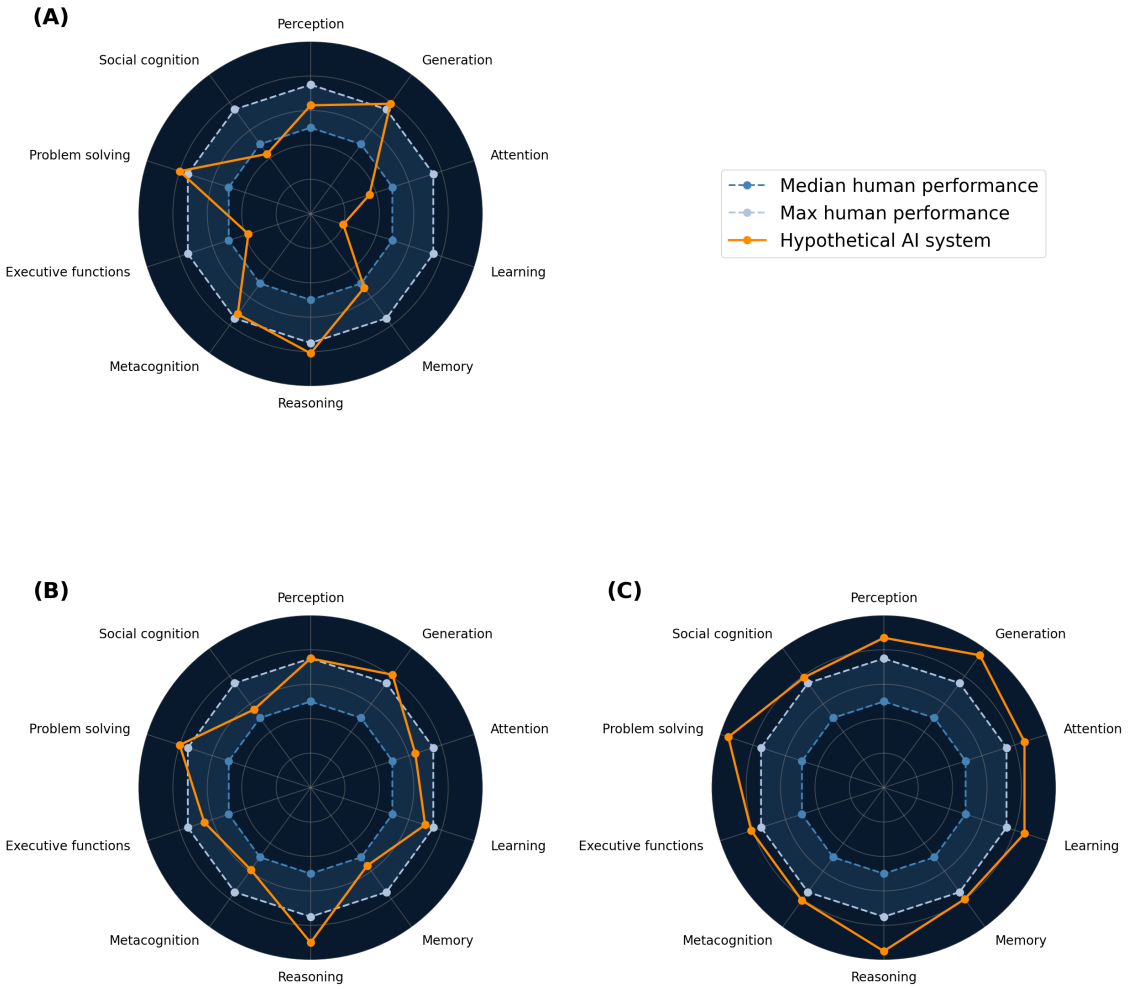


Figure 2 | *Cognitive profiles for three hypothetical systems. Panel A: A hypothetical system that shows significant cognitive weaknesses relative to the human sample. Panel B: A hypothetical system that outperforms the human sample median across all cognitive faculties. Panel C: A hypothetical system that outperforms the human sample maximum across all cognitive faculties. The hypothetical human sample scores are standardized for each dimension for illustrative purposes.*

4. Discussion

4.1. Building cognitive evaluations

Here we provide a framework for measuring AI systems across the range of human cognitive capabilities. To execute on this framework, we need robust cognitive benchmarks. Many useful benchmarks already exist for some parts of the cognitive space, including problem solving (Phan et al., 2025), perception (Patraucean et al., 2023), and world knowledge (Cheng et al., 2025; Haas et al., 2025). But there are large coverage gaps in areas such as metacognition, attention, learning, and social cognition. Moreover, many of the high-quality benchmarks that do exist are fully public, so they are susceptible to data contamination and may not provide generalizable signal. For this reason, we do not think existing benchmarks are sufficient to reliably evaluate AI cognitive capabilities.

To address these gaps, we are working with the academic community to build robust, held-out evaluations that will better enable us to evaluate increasingly capable AI systems.

4.2. Beyond cognitive faculties

Our taxonomy focuses on cognitive capabilities. But to build a full understanding of AI capabilities and AI behavior we will need to consider—and evaluate—many other system characteristics. Below we discuss several that we think will be important, all of which can be understood in relation to a human baseline.

4.2.1. Processing and response speed

For a response to be adaptive or useful, correctness is not always sufficient. Often, the response must be timely, too. Take, for example, a self-driving car system. The system's reliability depends not only on its ability to identify potential hazards, but its ability to do so quickly. Even when speed is less critical, it is still important for understanding real-world utility—a system that can fix a coding bug or book a flight in one minute is likely to be much more useful than one that takes six hours to complete the task.

The construct of processing speed is at least partly cognitive (Danthiir et al., 2005)—for example, problem solving speed depends on the thinking strategies a system employs and on how well it draws connections to relevant knowledge. But speed also depends on other factors such as hardware and network speed, so it does not fit perfectly as a cognitive faculty. Regardless, response speed is a key performance metric that should be measured across the cognitive space.

4.2.2. System propensities

Another important determinant of how a system will behave when deployed is its propensities (i.e. not just what the system *can* do but what it will *tend* to do). How willing is the system to take risks? How aligned is it with human values? What are its typical problem-solving strategies? How does it communicate and interact with people? These kinds of behavioral factors will significantly affect how safe and reliable that system is, so we need robust tools to evaluate them. A full account of system propensities is beyond the scope of this paper, but will be a critical area of study to inform deployment decisions and enable effective governance (see e.g. Romero-Alvarado et al. 2026; Taubenfeld et al. 2026).

4.2.3. Creativity

The human capacity for creativity has long been a topic of interest to both philosophy and cognitive science. Yet disagreements continue about how creativity should be conceptualized and measured (see e.g. Boden 1994; Kaufmann 2003). One common way of thinking about creative outputs is that they need to be both *novel* and *high-quality* (Sternberg and Lubart, 1998). But whether an output is high quality is highly subjective and domain specific, especially when it has entirely novel features. Moreover, it can be argued that these characteristics are simply general features of intelligent behavior. For this reason, it may be difficult to isolate and evaluate creativity objectively in AI systems. However, we can still evaluate the cognitive processes involved in creativity. For example, creativity is often associated with cognitive flexibility (the ability to switch modes of thinking and the ability to generate a wide variety of distinct ideas), which is covered in the taxonomy as part of executive functions. In addition, the taxonomy captures world knowledge (as part of semantic memory) and problem solving, both of which are relevant to creativity (Sternberg and Lubart, 1998).

4.2.4. End-to-end deployment evaluations

Lastly, it is important to make clear that cognitive benchmarking is not a substitute for applied, end-to-end evaluations—if a system is being deployed in a specific context, it is absolutely critical to evaluate the system on important deployment workflows. These two approaches are complementary—cognitive evaluations can help explain model failures and inform model improvements; real-world, deployment evaluations can inform deployment decisions and predict economic impacts (see, e.g. Mazeika et al. 2025; Patwardhan et al. 2025).

4.3. Model vs system evaluation

Historically, evaluations were largely focused on a specific model checkpoint. But modern AI systems are more than just a model—they are deployed with specific system instructions, have access to tools, can manipulate their environments via actions, and may even have the ability to make calls to other AI systems.

How should we approach evaluations in light of this added complexity? We think that attempting to isolate the core model of a system from its other components will become increasingly impractical since the inner workings of new systems are often not disclosed and these different components may not even be easily separable. This model-only approach will also become less and less informative, since the results would not be representative of how the system will operate or perform when deployed with access to these components. Modularity is also not unique to artificial systems—after all, the human brain has various cognitive systems and modules with different (but interconnected) functions (Yeo et al., 2011). We therefore think the best approach is to evaluate the system as a whole, including any built-in tools or modules.

Measuring intelligence as a property of the system does have its drawbacks, though. For example, one difficult ramification is that the intelligence attributed to a given model depends on the harness built around it—analogous to concluding that a person becomes more intelligent when given access to a calculator or a computer. In addition, this approach raises questions about how to construct informative cognitive tests. When testing humans, access to external tools is typically restricted to maintain control over the testing conditions and target specific cognitive processes. If we allow AI systems to have access to any and all tools during evaluations, these tools could muddy the interpretation of the findings. For instance, imagine we want to test semantic memory for historical events. If the system can simply search the internet for information about these events, we are no longer measuring the system's memory—only its ability to search the internet. These are thorny

issues that require consideration for each individual cognitive test. At the very least, any human baseline studies should ensure that participants are given access to the same tools we expect the AI systems to employ.

4.4. Validation and iteration

Like any science, the science of artificial general intelligence will be iterative in nature. The cognitive taxonomy described here is very much a starting point—it seems certain that AI systems will go on to develop cognitive capabilities that do not map neatly onto our taxonomy. Indeed, AI systems already possess some capabilities not found in humans, such as LiDAR perception and native image generation. Future iterations of this taxonomy will need to explore how to identify, characterize, and incorporate these emergent components of intelligence.

We also still have much to learn about how specific cognitive faculties are related to real-world performance. If a system lags behind humans in one or more of the 10 faculties, it would clearly demonstrate the system cannot match the generality of human intelligence. But what would a weakness in a given faculty mean when the system is deployed in the real world? There are good theoretical reasons to expect that each one of the 10 cognitive capacities identified here is important for different aspects of real-world performance. For example, a system that cannot plan would likely struggle with long-horizon, multi-step tasks, while a system deficient in social understanding is likely to perform poorly in situations that involve complex interactions with people. But empirical work is still needed to demonstrate these relationships and to further understand the importance of each cognitive capacity when it comes to practical, real-world tasks.

5. Conclusion

The pursuit of Artificial General Intelligence represents a pivotal moment for humanity. By providing a clear, empirical framework rooted in the established science of human cognition, we aim to move the conversation around AGI from one of subjective claims and speculation toward a grounded, measurable scientific endeavor. Our Cognitive Taxonomy and evaluation protocol offer a way to map the jagged landscape of AI capabilities and track progress toward general intelligence.

6. Contributors and Acknowledgments

Contributors

Alex Siegman
Alëna Aksënova
Anastasios Kementsietsidis
Arjun Narayanan
Ashwin Vaswani
Bernd Bohnet
Chrysovalantis Anastasiou
Claire Yao
Daniel McDuff
Dima Yeroshenko
Edward Loper
Ellie Pavlick
Evan Rosen
Georgi Karadzhov
Jed McGiffin
Joe Heyward
Julia Haas
Kartikeya Badola
Kate Lin
Laura Kampis
Martin Polacek
Martyna Płomecka
Mor Hazan Taege
Nicholas Cain
Nico Duduta
Rasmi Elasmar
Reut Aharony
Rivka Moroshko
Silvia Chiappa
Steve Zheng
Vik Sharma
Viorica Patraucean
Virginia Aglietti
William Cunningham
William Isaac
Xin Liu
Yao Yan
Yuan Yuan

This effort involved the contributions of many individuals across Google, including researchers, engineers, and operations staff. We gratefully acknowledge the dedication and hard work of each contributor on this effort. Contributors are listed in alphabetical order.

We are grateful for the invaluable feedback from Dharshan Kumaran, Joel Z Leibo, Ellie Pavlick, Mike Mozer, Martin Chadwick, Zoubin Ghahramani, Rohin Shah, Iason Gabriel, Lucy Cheke, and Jose Hernandez-Orallo.

7. Appendix: Cognitive taxonomy

7.1. Perception

The ability to take in and process sensory information from the world such as images, audio, and text (Harris and Smith, 2022). Perception allows a system to observe the environment and respond to its characteristics.[3] We organize this section by modality, and include aspects of language processing in the relevant modality.

7.1.1. Visual perception

The ability to take in and process visual information from light (Cornsweet, 1970; Haber and Hershenson, 1973).

Low-level visual perception The ability to detect and identify surface-level features of visual information, such as light intensity, color, and contrast. [4]

- **Light detection:** The ability to detect the brightness of different parts of a visual scene (Gilchrist et al., 1999).
- **Contrast (edge) detection:** The ability to detect contrast between light and dark areas of an image (Marr and Hildreth, 1980). Critical for image segmentation.
- **Color detection:** The ability to distinguish between different colors in an image (Gegenfurtner, 2003).
- **Depth perception:** The ability to identify the distance to objects in a visual scene (Parker, 2007).
- **Motion detection:** The ability to detect movement in a dynamic visual scene (Borst and Egelhaaf, 1989).
- **Shape detection:** The ability to identify and locate simple shapes in an image (Biederman, 1987; Todd, 2004).

High-level visual perception The ability to process, interpret, and understand the semantic meaning of visual information (Ullman, 1996).

- **Object recognition:** The ability to identify and categorize objects in a visual scene (Riesenhuber and Poggio, 2000).
- **Visual spatial localization:** The ability to determine the spatial location of objects in a visual scene (Hollingworth, 2007).
- **Visual counting:** The ability to determine the number of objects or entities in a visual scene (Trick and Pylyshyn, 1994).
- **Static scene understanding (image understanding):** The ability to form a high-level understanding of events appearing in a static visual scene (Epstein and Baker, 2019).
- **Dynamic scene understanding (video understanding):** The ability to form a high-level understanding of events occurring in a dynamic visual scene (Zacks et al., 2010).

[3] Input modalities are highly system dependent. A system may not need to perceive **all** modalities to behave intelligently, although each additional modality of information is likely to provide some benefits.

[4] Here, we classify abilities as low-level perception when they involve "surface-level" processing with little semantic interpretation. Conversely, we classify abilities that require deeper, semantic processing as high-level visual perception. This is because there is evidence in humans that high-level and low-level perceptual processing can be dissociated (e.g. Farah 1990). However, the dividing line is somewhat fuzzy—perceptual abilities fall on a spectrum from those which require very little semantic processing (e.g., contrast detection) to those which are highly reliant on semantic processing (e.g., scene understanding).

- **Word segmentation:** The ability to segment and identify letters and words in a visual scene (Norris, 2013).
- **Reading comprehension:** The ability to understand the meaning of language presented in a visual form (Kendeou et al., 2016).[5]

7.1.2. Auditory perception

The ability to take in and process auditory information from sound waves (Warren, 1982).

Low-level Auditory perception The ability to extract surface-level features of auditory information such as loudness, pitch, and spatial location.

- **Loudness detection:** The ability to identify the loudness (volume) of a sound (Fechner, 1860).
- **Pitch detection:** The ability to identify the pitch of a sound (Oxenham, 2012).
- **Sound discrimination:** The ability to differentiate components of an auditory scene occurring simultaneously, such as different sounds or speakers (Bizley and Cohen, 2013).
- **Sound localization:** The ability to identify the direction of a sound and the distance to its source (Blauert, 1996).
- **Speech segmentation:** The ability to separate a stream of sound information into separate words (Brent, 1999).
- **Rhythm perception:** The ability to identify patterns or rhythms in sound information (Grahn, 2012).

High-level Auditory perception The ability to process, interpret, and understand auditory information.

- **Speech comprehension:** The ability to understand the meaning of spoken language (Diehl et al., 2004).
- **Speaker identification:** The ability to recognize and distinguish individual speakers (McDermott, 2009).
- **Sound recognition:** The ability to recognize different types of sounds.
- **Auditory scene understanding:** The ability to understand and interpret events occurring in an auditory scene (Bregman, 1994).
- **Listening comprehension:** The ability to understand the meaning of language presented in auditory form (Friederici, 2002).[6]

7.1.3. Text perception

The ability to process, interpret, and understand information presented as text.

Unlike humans, who perceive text only through other modalities (e.g. through vision via reading, or through touch via braille reading), today's AI systems perceive text as an entirely separate modality via tokenization and embedding. These processes are flawed and lossy much like other forms of perceptual processing (Shin and Kaneko, 2024)—as illustrated by systems' difficulties with spelling (how many r's are in strawberry, again?) and their brittleness to perturbations in text inputs.

[5] To demonstrate reading comprehension, a system needs to show mastery of at least one language. The ability to learn new languages is covered under learning, and the number of languages a system knows is a function of its semantic memory.

[6] As with reading comprehension, to demonstrate listening comprehension a system needs to show mastery of at least one language. The ability to learn new languages is covered under learning, and the number of languages a system knows is a function of its world knowledge.

Although text perception is not strictly a fundamental capability that humans have, textual information is central to today's society and is an essential part of how today's systems are trained to understand language and the world. The ability to bypass vision and directly perceive text-based language is a fascinating and unique property of today's AI systems. Language models have demonstrated that this ability can enable remarkable performance on a range of tasks, so we think it is worthy of consideration and measurement. Of course, the lack of a direct equivalent in humans raises questions about how text perception should factor into judgments about AGI. In one sense, this ability can be thought of as an "easier" version of reading, without the need to segment the visual information into letters and words. Therefore, we think that an AGI should, at a minimum, be able to comprehend a given text at least as well as a human could if presented with that text in visual form.

Low-level text perception The ability to extract surface-level features from text inputs such as letters and words.

- **Symbol discrimination:** The ability to discriminate between different symbols in a stream of characters (e.g. letters and numbers).
- **Word segmentation:** The ability to segment a stream of characters into words.

High-level text perception The ability to process, interpret, and understand text information.

- **Language comprehension:** The ability to understand the meaning of language presented as text (Traxler and Gernsbacher, 2006).
- **Code comprehension:** The ability to understand the meaning and function of written code.

7.1.4. Other modalities of perception

There are many other types of perception seen in humans or other animals that may provide utility for AI systems. It seems likely that each additional modality of perception will provide some useful information to systems that can help them navigate the real world. For example, touch (de Haan and Dijkerman, 2020), smell (Stevenson, 2009), and temperature perception (Jung et al., 2023) could all prove highly useful for a system designed to help with cooking or manufacturing. Still, the relative importance of these modalities for intelligent behavior remains unclear, so in this initial version of the taxonomy we do not include them in detail. These modalities are nonetheless worth considering, both for researchers building new systems and those building model evaluations.

7.1.5. Multi-sensory integration

The ability to integrate information from multiple modalities together. Because different perceptual modalities provide complementary information about features such of the environment, such as form, location, size, and speed of movement, multi-sensory integration helps enable joint processing, reasoning, and planning across the information (Zmigrod and Hommel, 2013). This ability should be measured for each combination of modalities.

7.2. Generation

The ability to generate outputs such as text, speech, or actions such as motor movements, computer control actions, or tool use calls. Output generation is critical for a system's ability to communicate and complete tasks effectively.

The ability to generate high-quality outputs can be at least partly decoupled from the ability to decide which output to generate (imagine a very savvy but poorly-skilled tennis player who might correctly determine the best shot to play but fail miserably to execute that shot effectively). In other words, output generation captures a system's *execution* ability, not its ability to reason and plan about which actions to attempt.

7.2.1. Text generation

The ability to generate text outputs. In humans, text generation is indirect, typically mediated through motor actions (via writing or typing), but in AI systems text can be a direct output.

Natural language generation The ability to generate natural language in text form. Typically known as language *production* in cognitive science (Garrett, 1988; Pickering and Garrod, 2013).

- **Grammatical correctness:** The ability to produce grammatically correct sentences.
- **Lexical selection:** The ability to select appropriate words to convey rich meaning.

Code generation The ability to generate structurally and syntactically correct code (see Fedorenko et al. 2019 for a discussion).

7.2.2. Audio generation

The ability to produce audio outputs (i.e., sound).

Speech generation The ability to generate natural and expressive speech (Traxler and Gernsbacher, 2006).

- **Clarity:** The ability to clearly pronounce words and phonemes.
- **Grammatical correctness:** The ability to produce grammatically correct sentences.
- **Lexical selection:** The ability to select appropriate words to convey rich meaning.
- **Prosody control:** The ability to control the rhythm of speech such as pitch, stress, speed and intonation.
- **Emotional expression:** The ability to express a wide range of emotions through variation in prosody.

7.2.3. Action generation

The ability to generate actions that manipulate the environment.

Motor control The ability to control a body or robotic actuators (Nishikawa et al., 2007).

Computer control The ability to produce computer control actions such as key presses and mouse movements (Smith et al., 1999).

Tool use The ability to use external objects, systems, or resources to assist in the pursuit of goals (Baber, 2003).

7.2.4. Thought generation

The ability to generate internal thoughts which can be used to guide decisions (Holyoak and Spellman, 1993). Thoughts could take the form of language, images (akin to human visual imagination), or could take a more abstract form.

By definition thoughts are internal in nature and may be difficult or impossible to evaluate. However, conscious thought is critical for human problem solving and there is substantial evidence for its value in AI systems (Comanici et al., 2025; OpenAI, 2024), so evaluating the features and quality of a system's thoughts will almost certainly be enlightening.

7.3. Attention

The ability to focus cognitive resources on specific aspects of perceptual stimuli, information, or thoughts (Styles, 2006; Treisman, 1969). This is critical when a system is faced with a complex environment and has limited cognitive resources.

There is a delicate balance between the need to avoid distraction by narrowly focusing on information that is important for current goals and the need to stay attentive to the wider environment to respond when unexpected changes or stimuli appear (Burgoyne and Engle, 2020; Engle, 2018).

7.3.1. Attention capacity

The amount of information a system can focus on simultaneously. This capacity may be different for different modalities or types of information (e.g., text, images, audio; Cowan et al. 2005; Fritz et al. 2007).

7.3.2. Selective attention / Attentional control

The ability to selectively focus on information that is relevant to current goals and ignore information that is goal-irrelevant (Theeuwes, 1991). This active, top-down form of attention is important for effective goal-driven behavior and is often considered an executive function.

Sustained attention The ability to maintain focus on goal-relevant information over time (Esterman and Rothlein, 2019; Sarter et al., 2001).

Perceptual inhibition The ability to ignore distracting or goal-irrelevant perceptual information (Van Moorselaar and Slagter, 2020).

Attention shifting The ability to actively shift attention from one location or piece of information to another (Brown and Tait, 2016).

7.3.3. Stimulus-driven attention

The ability for attention to be directed in a "bottom-up" way toward new stimuli or environmental changes (Awh et al., 2012; Katsuki and Constantinidis, 2014). This is crucial for quickly identifying situational shifts that require a response.

7.4. Learning

The ability to acquire new knowledge, skills, or behaviors through experience, study, or instruction. Learning is vital for a system's ability to adapt to new situations or environmental changes (Ginsburg and Jablonka, 2010; Morand-Ferron, 2017).[7] Here, we enumerate several important kinds of learning an AGI should be able to employ.

[7]For many current systems, learning occurs only during training or in-context. However, for truly robust and adaptive behavior, AI systems should be able to learn (and retain) new knowledge and skills over time (e.g., as part of a continuous learning process).

7.4.1. Concept formation

The ability to abstract the key features of objects, events, and ideas to form categories, concepts, schemas, and scripts (Bruner, 1986; Gershman and Niv, 2010; Goodman et al., 2008; Tenenbaum et al., 2011). Important for generalization.

7.4.2. Associative learning

The ability to learn the relationships between events, objects or stimuli that appear together (Shanks, 1995).

7.4.3. Reinforcement learning (Operant conditioning)

The ability to learn based on the consequences (rewards and punishments) of specific actions or situations (Skinner, 1963; Sutton et al., 1998).

7.4.4. Observational learning

The ability to learn by watching others perform a skill or task (Bandura, 2008).

7.4.5. Procedural learning

The ability to learn skills, action patterns, or tasks through performance or practice (Cohen and Squire, 1980). [8]

7.4.6. Language learning

The ability to learn new language-related information, such as syntax and vocabulary (Bates, 1976; Chomsky, 1965; Pinker, 1994; Saffran, 2003). Includes natural languages as well as coding languages and tool use frameworks.

7.5. Memory

The ability to keep track of information over time (Squire and Kandel, 1999). Memory and learning are closely linked—the distinction is that learning is focused on the acquisition of new knowledge, whereas memory is concerned with the ability to maintain that knowledge over time. Evaluating memory typically involves testing a system's pre-existing knowledge as well as its ability to store and retrieve newly-learned information. The quality of a system's memory can be considered in terms of how much information a system can remember, how long the information can be maintained, and/or how accurate the memories are.[9]

Given how closely linked learning and memory are, it may be difficult to separate them when evaluating system capabilities. However, there are at least some failure modes that are specific to one or the other. For example, a failure to update semantic knowledge despite being able to successfully recall already stored knowledge would be considered a failure of learning, while forgetting information over time that was initially successfully learned would be a failure of memory.

[8] In humans, procedural learning can be dissociated from the ability to learn facts and other explicit information (Willingham et al., 1989).

[9] Most research on human memory is focused on understanding the specific memory mechanisms of the brain, such as "long-term memory" and "short-term memory". But these mechanisms are specific to humans and may not be relevant to AI systems. For this reason, we think it is important to remain agnostic to the underlying memory implementation and focus on evaluating *how well* a system can keep track of information over time.

7.5.1. Semantic memory (World knowledge)

The ability to keep track of facts and other general information not tied to a specific episode (Kumar, 2021; Yee et al., 2013).

Commonsense knowledge Knowledge about the fundamental rules, properties, and characteristics of the world. Includes:

- General knowledge
- Causal knowledge
- Temporal knowledge
- Spatial knowledge
- Intuitive physics knowledge

Domain knowledge Specialist knowledge about specific subjects or domains (Ericsson et al., 2018). For example:

- Linguistic knowledge
- STEM knowledge
- Coding knowledge
- Legal knowledge
- Financial knowledge
- Medical knowledge
- Historical knowledge
- Social-cultural knowledge

7.5.2. Episodic memory

The ability to keep track of information relating to specific events, including the sensory information (images, sounds, etc.) associated with those events (Conway, 2009; Tulving et al., 1972).[10]

Sensory memory The ability to keep track of sensory information from specific events (e.g. visual, auditory, or olfactory information).

Temporal memory The ability to keep track of the sequence and temporal relationships between events.

Spatial memory The ability to keep track of the spatial relationships between objects from specific events (e.g. their positions, orientations, and movements).

7.5.3. Procedural memory

The ability to keep track of the action and output patterns needed to perform skills (Cohen and Squire, 1980).

7.5.4. Prospective memory

The ability to remember to perform a planned action when a specific cue arises, such as a moment in time or in response to a specific event (e.g. “When I go to the store, I should remember to get milk”;

[10]The ability to remember sensory information is sometimes separated from the ability to remember factual details about specific events. The relationship between these aspects of event memory is complex, so for practical purposes we use episodic memory to refer to both together.

McDaniel and Einstein 2007).

7.5.5. Forgetting

The ability to remove outdated, wrong, or irrelevant information from memory. This could involve information compression or wholesale pruning of specific memories. Forgetting is important for efficient storage and retrieval (Bjork, 1989).

7.6. Reasoning

The ability to draw valid conclusions by applying logical principles or drawing inferences. (Leighton and Sternberg, 2003; Rips, 1990). Typically, automatic pattern matching would not be considered reasoning.

7.6.1. Deductive (logical) reasoning

The ability to reason from a set of premises, rules or facts to reach a logical conclusion (Johnson-Laird, 1999). A defining feature of deduction is a lack of ambiguity—as long as one can be sure the premises are true, then correctly applied deduction will lead to conclusions that are certain. Deduction requires a grasp of logical concepts such as negation, "AND", "OR", "XOR", and "ALL".

7.6.2. Inductive reasoning

The ability to draw general conclusions based on a specific set of facts, information, or observations (e.g., "The sun has risen every day so far; therefore, the sun will rise tomorrow").

Unlike deductive reasoning, these general conclusions are probabilistic, not certain. As a result, induction can occasionally lead us astray. However, in a world in which there is constant uncertainty, inductive reasoning is essential to avoid constant paralysis (Heit, 2000).

7.6.3. Abductive reasoning

The ability to make inferences about the best or most likely explanation for a set of observations. Much like inductive reasoning, abductive reasoning is uncertain. A key feature of abduction is that it involves the generation of new explanatory hypotheses, rather than just generalizing from a set of observations (Bhagavatula et al., 2020).

7.6.4. Analogical reasoning

The ability to identify similarities between situations or concepts and to use those similarities to draw conclusions about unknown properties of one based on known information about the other (Alexander et al., 2016; Gentner and Maravilla, 2018).

7.6.5. Mathematical reasoning

The ability to perform mathematical calculations and operations (Gilmore et al., 2018). Includes basic arithmetic and algebra.

7.7. Metacognition

The knowledge a system has about its own cognitive processes, and its ability to monitor and control those processes (Dunlosky and Metcalfe, 2009; Tarricone, 2011).

7.7.1. Metacognitive knowledge

Metacognitive knowledge is a system's self-knowledge about its own abilities, limitations, knowledge, learning processes, and behavioral tendencies (Flavell, 1979; Tarricone, 2011). In some ways, metacognitive knowledge could simply be considered a special case of world knowledge.

Knowledge of limitations Knowledge of one's own capabilities and limitations (Fleming and Daw, 2017; Fleming and Lau, 2014).

Knowledge of learning processes Knowledge of one's own learning processes and the factors that can assist or impair learning (Binz et al., 2024; Wang, 2021).

Metamemory (Knowledge of knowledge) Knowledge about the information stored in memory and about the processes involved in storing or retrieving that information (Nelson, 1990).

Knowledge of behavioral patterns Knowledge of one's own tendencies and behavioral patterns (Grant, 2001; Vazire and Carlson, 2010).

7.7.2. Metacognitive monitoring

The ability to monitor the state of cognitive processes (e.g., evaluating the state of learning or current performance) (Dunlosky and Metcalfe, 2009; Nelson, 1990).

Confidence calibration The ability to accurately estimate the likelihood that an action will be successful or that a response will be correct (Fleming and Lau, 2014; Harvey, 1997; Yeung and Summerfield, 2012)

Judgments of learning The ability to monitor progress when learning new information (Arbuckle and Cuddy, 1969; Rhodes, 2016)

Error monitoring The ability to notice when errors are made (Yeung and Summerfield, 2012).

Source judgments The ability to judge where a piece of information was generated or learned from (Johnson et al., 1993; Mitchell and Johnson, 2000)

7.7.3. Metacognitive control

A model's ability to utilize insights from metacognitive knowledge and monitoring to adjust cognitive processes or strategies (e.g. by switching learning strategy based on the kinds of errors that one is making; Botvinick 2007; Nelson 1990; Son and Schwartz 2002). Sometimes considered an executive function.

Error correction The ability to adjust action patterns or strategies to correct errors (Metcalfe, 2017).

Learning strategy selection The ability to select appropriate learning strategies based on metamemory and judgments of learning (e.g. terminating study of well-learned information to focus on information that has not yet been well-learned) (Dunlosky et al., 2013).

7.8. Executive functions

Higher-order cognitive abilities that enable goal-directed behavior by regulating and orchestrating thoughts and actions (Diamond, 2013).

7.8.1. Goal setting and maintenance

The ability to set and maintain goals to organize and direct action (Buschman and Miller, 2014; Dickinson and Balleine, 1994).

7.8.2. Planning

The ability to formulate sequences of future actions to achieve specific goals (Owen, 1997). Planning is a key part of solving multi-step or long-term problems. The process of planning can be broadly construed as building up (and pruning) some kind of decision-tree (Mattar and Lengyel, 2022).

7.8.3. Inhibitory control

The ability to change, withhold, or suppress learned or habitual responses in favor of more controlled, goal-appropriate ones (Bari and Robbins, 2013; Miyake et al., 2000).

7.8.4. Cognitive flexibility

The ability to switch between different tasks, concepts, or ways of thinking (Braem and Egner, 2018).

7.8.5. Conflict resolution

The ability to manage and resolve conflicting information, contradictory goals, or competing responses to select an appropriate action (Botvinick et al., 2001; Veen and Carter, 2006). Not to be confused with social conflict resolution.

7.8.6. Working memory

The ability to manipulate information internally in service of a goal (e.g. performing intermediate calculations while solving a problem or mentally rotating an image to consider it from a different perspective) (Baddeley, 1992). Although the name would suggest this ability is a subset of memory, in truth working memory involves the coordination of multiple faculties including memory, attention, and sometimes reasoning (Engle, 2002).

7.9. Problem solving

As the name suggests, this broad ability refers to the ability to solve problems and overcome obstacles (Mayer and Wittrock, 2006). This is a composite ability that relies heavily on planning, reasoning, and in-context learning. Problem solving requires:

- Understanding and representing the problem (e.g. via perception and abstraction)
- Identifying and retrieving relevant knowledge (e.g., facts, analogous episodes, meta-knowledge about effective problem-solving strategies).
- Breaking down the problem into sub-goals
- Planning a sequence of actions to take

- Executing the plan (e.g. via reasoning + output generation)

It would be impossible to enumerate all types of problems that humans are able to solve, but we describe several important types below.

7.9.1. Fluid reasoning

The ability to identify patterns and apply them to solve novel problems. Relies on a mix of deductive, inductive, and abductive reasoning (Cattell, 1943; Kent, 2017).

7.9.2. Mathematical problem solving

Applying mathematical concepts and techniques to solve problems (Schoenfeld, 1985).

7.9.3. Algorithmic problem solving

Solving logical problems using algorithmic techniques—a key part of writing code.

7.9.4. Commonsense problem solving

Solving real-world problems by applying general knowledge and everyday understanding of the world (Brachman and Levesque, 2022).

Temporal problem solving Understanding and reasoning about temporal concepts such as the order and duration of events, time-based relationships, and temporal constraints.

Spatial problem solving Reasoning about the relationships between objects in space, including their positions, orientations, and movements.

Causal problem solving Reasoning about cause-and-effect relationships between events or entities.

Intuitive physics Reasoning about basic physical principles and how objects behave in the world. Includes concepts such as:

- Object permanence
- Gravity
- Momentum
- Force

7.9.5. Knowledge discovery

The ability to generate novel hypotheses, experiments, and solutions to scientific questions (Dunbar, 2001; Klahr, 2000; Nersessian, 2002).

7.10. Social cognition

The ability to process and interpret social information and to respond appropriately in social situations. Crucial for interacting with people or other systems (Higgins and Bargh, 1987).

7.10.1. Social perception

The ability to interpret social cues based on perceptual information such as facial expressions, tone of voice, or body language (Tajfel, 1962).

7.10.2. Theory of mind

The ability to reason about the mental states of others, including beliefs, desires, emotions, intentions, expectations, and perspectives. Theory of mind is important for the ability to predict and explain others' behavior (Frith and Frith, 2005; Leslie et al., 2004).

7.10.3. Social skills

The ability to recognize, understand, and act according to social norms or expectations (Chung and Rimal, 2016).

Cooperation The ability to work together with others toward common goals (Rand and Nowak, 2013).

Negotiation The ability to work together with others toward goals that are misaligned or in conflict. (Bazerman et al., 2000).

Deception The ability to mislead others by hiding or disguising intentions or actions in order to achieve goals (Spence et al., 2004). Could be considered a harmful capability depending on the context.

Persuasion The ability to influence others' attitudes, beliefs, or behaviors (Wood, 2000). Could be considered a harmful capability depending on the context.

References

P. A. Alexander, D. Dumas, E. M. Grossnickle, A. List, and C. M. Firetto. Measuring relational reasoning. *The Journal of Experimental Education*, 84(1):119–151, 2016. doi: 10.1080/00220973.2014.963216. URL https://doi.org/10.1080/00220973.2014.963216.

T. Y. Arbuckle and L. L. Cuddy. Discrimination of item strength at time of presentation. *Journal of experimental psychology*, 81(1):126, 1969.

E. Awh, A. V. Belopolsky, and J. Theeuwes. Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in cognitive sciences*, 16(8):437–443, 2012.

C. Baber. *Cognition and tool use: Forms of engagement in human and animal use of tools*. CRC Press, 2003.

A. Baddeley. Working memory. *Science*, 255(5044):556, 1992.

A. Bandura. Observational learning. In *The international encyclopedia of communication*. Wiley Online Library, 2008.

A. Bari and T. W. Robbins. Inhibition and impulsivity: behavioral and neural basis of response control. *Progress in neurobiology*, 108:44–79, 2013.

E. Bates. *Language and context: The acquisition of pragmatics*. Academic Press, 1976.

M. H. Bazerman, J. R. Curhan, D. A. Moore, and K. L. Valley. Negotiation. *Annual review of psychology*, 51(1):279–314, 2000.

C. Bhagavatula, R. L. Bras, C. Malaviya, K. Sakaguchi, A. Holtzman, H. Rashkin, D. Downey, S. W.-t. Yih, and Y. Choi. Abductive commonsense reasoning, 2020. URL https://arxiv.org/abs/1908.05739.

I. Biederman. Recognition-by-components: a theory of human image understanding. *Psychological review*, 94(2):115, 1987.

M. Binz, I. Dasgupta, A. K. Jagadish, M. Botvinick, J. X. Wang, and E. Schulz. Meta-learned models of cognition. *Behavioral and Brain Sciences*, 47:e147, 2024.

J. K. Bizley and Y. E. Cohen. The what, where and how of auditory-object perception. *Nature Reviews Neuroscience*, 14(10):693–707, Oct 2013. ISSN 1471-0048. doi: 10.1038/nrn3565. URL https://doi.org/10.1038/nrn3565.

R. A. Bjork. *Retrieval inhibition as an adaptive mechanism in human memory*, pages 309–330. Varieties of memory and consciousness: Essays in honour of Endel Tulving. Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, US, 1989. ISBN 0-89859-935-0 (Hardcover); 0-8058-0546-X (Paperback).

J. Blauert. *Spatial Hearing: The Psychophysics of Human Sound Localization*. The MIT Press, 10 1996. ISBN 9780262268684. doi: 10.7551/mitpress/6391.001.0001. URL https://doi.org/10.7551/mitpress/6391.001.0001.

M. A. Boden. What is creativity? In M. A. Boden, editor, *Dimensions of Creativity*, pages 75–117. The MIT Press, Cambridge, MA, 1994. ISBN 9780262023689. doi: 10.7551/mitpress/2437.003.0006.

A. Borst and M. Egelhaaf. Principles of visual motion detection. *Trends in neurosciences*, 12(8): 297–306, 1989.

M. M. Botvinick. Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4):356–366, 2007.

M. M. Botvinick, T. S. Braver, D. M. Barch, C. S. Carter, and J. D. Cohen. Conflict monitoring and cognitive control. *Psychological review*, 108(3):624, 2001.

R. J. Brachman and H. J. Levesque. *Machines like Us: Toward AI with Common Sense*. The MIT Press, Cambridge, MA, 2022. ISBN 9780262369237. doi: 10.7551/mitpress/14299.001.0001. URL https://doi.org/10.7551/mitpress/14299.001.0001.

S. Braem and T. Egner. Getting a grip on cognitive flexibility. *Current directions in psychological science*, 27(6):470–476, 2018.

A. S. Bregman. *Auditory scene analysis: The perceptual organization of sound*. MIT press, 1994.

M. R. Brent. Speech segmentation and word discovery: A computational perspective. *Trends in Cognitive Sciences*, 3(8):294–301, 1999.

V. J. Brown and D. S. Tait. *Attentional Set-Shifting Across Species*, volume 28 of *Current Topics in Behavioral Neurosciences*, pages 363–395. Springer International Publishing, Cham, 2016. ISBN 978-3-319-33913-9. doi: 10.1007/7854_2015_5002. URL https://doi.org/10.1007/7854_2015_5002.

J. S. Bruner. *A Study of Thinking*. Routledge, 2nd edition, 1986. doi: 10.4324/9781315083223.

A. P. Burgoyne and R. W. Engle. Attention control: A cornerstone of higher-order cognition. *Current Directions in Psychological Science*, 29(6):624–630, 2020.

T. J. Buschman and E. K. Miller. Goal-direction and top-down control. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1655):20130471, 2014.

R. B. Cattell. The measurement of adult intelligence. *Psychological bulletin*, 40(3):153, 1943.

A. Cheng, A. Jacovi, A. Globerson, B. Golan, C. Kwong, C. Alberti, C. Tao, E. Ben-David, G. S. Tomar, L. Haas, Y. Bitton, A. Bloniarz, A. Bai, A. Wang, A. Siddiqui, A. B. Castillo, A. Atias, C. Liu, C. Fry, D. Balle, D. Ghosal, D. Kukliansky, D. Marcus, E. Gribovskaya, E. Ofek, H. Zhuang, I. Laish, J. Ackermann, L. Wang, M. Risdal, M. Barnes, M. Fink, M. Amin, M. Ambar, N. Potikha, N. Gupta, N. Katz, N. Velan, O. Roval, O. Ram, P. Zablotskaia, P. Bang, P. Agrawal, R. Ghiya, S. Ganapathy, S. Baumgartner, S. Erell, S. Prakash, T. Sellam, V. Rao, X. Wang, Y. Akulov, Y. Yang, Z. Yang, Z. Lai, Z. Wu, A. Dragan, A. Hassidim, F. Pereira, S. Petrov, S. Venkatachary, T. Doshi, Y. Matias, S. Goldshtein, and D. Das. The facts leaderboard: A comprehensive benchmark for large language model factuality, 2025. URL https://arxiv.org/abs/2512.10791.

P. W. Cheng and K. J. Holyoak. Pragmatic reasoning schemas. *Cognitive Psychology*, 17(4):391–416, 1985. ISSN 0010-0285. doi: https://doi.org/10.1016/0010-0285(85)90014-3. URL https://www.sciencedirect.com/science/article/pii/0010028585900143.

F. Chollet. On the measure of intelligence. *arXiv preprint arXiv:1911.01547*, 2019.

F. Chollet, M. Knoop, G. Kamradt, B. Landers, and H. Pinkard. Arc-agi-2: A new challenge for frontier ai reasoning systems. *arXiv preprint arXiv:2505.11831*, 2025.

N. Chomsky. *Aspects of the theory of syntax*. MIT Press, 1965.

A. Chung and R. N. Rimal. Social norms: a review. *Review of Communication Research*, 4:1–28, 2016. ISSN 2255-4165. doi: https://doi.org/10.12840/issn.2255-4165.2016.04.01.008.

N. J. Cohen and L. R. Squire. Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, 210(4466):207–210, 1980.

G. Comanici, E. Bieber, M. Schaekermann, I. Pasupat, N. Sachdeva, I. Dhillon, M. Blistein, O. Ram, D. Zhang, E. Rosen, et al. Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities. *arXiv preprint arXiv:2507.06261*, 2025.

M. A. Conway. Episodic memories. *Neuropsychologia*, 47(11):2305–2313, 2009.

T. N. Cornsweet. *Visual Perception*. Academic Press, 1970. ISBN 978-0-12-189750-5.

N. Cowan, E. M. Elliott, J. S. Saults, C. C. Morey, S. Mattox, A. Hismjatullina, and A. R. Conway. On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive psychology*, 51(1):42–100, 2005.

V. Danthiir, R. D. Roberts, R. Schulze, and O. Wilhelm. Mental speed: On frameworks, paradigms, and a platform for the future. In O. Wilhelm and R. W. Engle, editors, *Handbook of Understanding and Measuring Intelligence*, pages 27–46. Sage Publications, Inc., Thousand Oaks, CA, 2005. doi: 10.4135/9781452233529.n3.

I. Dasgupta, A. K. Lampinen, S. C. Y. Chan, H. R. Sheahan, A. Creswell, D. Kumaran, J. L. McClelland, and F. Hill. Language models show human-like content effects on reasoning tasks, 2024. URL https://arxiv.org/abs/2207.07051.

E. H. de Haan and H. C. Dijkerman. Somatosensation in the brain: a theoretical re-evaluation and a new model. *Trends in Cognitive Sciences*, 24(7):529–541, 2020.

A. Diamond. Executive functions. *Annual review of psychology*, 64(1):135–168, 2013.

A. Dickinson and B. Balleine. Motivational control of goal-directed action. *Animal learning & behavior*, 22(1):1–18, 1994.

R. L. Diehl, A. J. Lotto, and L. L. Holt. Speech perception. *Annu. Rev. Psychol.*, 55(1):149–179, 2004.

K. Dunbar. What scientific thinking reveals about the nature of cognition. In K. Crowley, C. D. Schunn, and T. Okada, editors, *Designing for science: Implications from everyday, classroom, and professional settings*, pages 115–140. Lawrence Erlbaum Associates, 2001.

J. Dunlosky and J. Metcalfe. *Metacognition*. Metacognition. Sage Publications, Inc, Thousand Oaks, CA, US, 2009. ISBN 978-1-4129-3972-0 (Paperback).

J. Dunlosky, K. A. Rawson, E. J. Marsh, M. J. Nathan, and D. T. Willingham. Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public interest*, 14(1):4–58, 2013.

R. W. Engle. Working memory capacity as executive attention. *Current directions in psychological science*, 11(1):19–23, 2002.

R. W. Engle. Working memory and executive attention: A revisit. *Perspectives on psychological science*, 13(2):190–193, 2018.

R. A. Epstein and C. I. Baker. Scene perception in the human brain. *Annual review of vision science*, 5 (1):373–397, 2019.

K. A. Ericsson, R. R. Hoffman, A. Kozbelt, and A. M. Williams, editors. *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge Handbooks in Psychology. Cambridge University Press, 2 edition, 2018.

M. Esterman and D. Rothlein. Models of sustained attention. *Current opinion in psychology*, 29: 174–180, 2019.

M. J. Farah. *Visual agnosia: disorders of object recognition and what they tell us about normal vision*. The MIT Press, 1990.

G. T. Fechner. *Elemente der psychophysik*, volume 2. Breitkopf u. Härtel, 1860.

E. Fedorenko, A. Ivanova, R. Dhamala, and M. U. Bers. The language of programming: A cognitive perspective. *Trends in Cognitive Sciences*, 23(7):525–528, 2019. ISSN 1364-6613. doi: https://doi.org/10.1016/j.tics.2019.04.010. URL https://www.sciencedirect.com/science/article/pii/S1364661319301020.

J. H. Flavell. Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, 34(10):906, 1979.

S. M. Fleming and N. D. Daw. Self-evaluation of decision-making: A general bayesian framework for metacognitive computation. *Psychological review*, 124(1):91, 2017.

S. M. Fleming and H. C. Lau. How to measure metacognition. *Frontiers in human neuroscience*, 8:443, 2014.

A. D. Friederici. Towards a neural basis of auditory sentence processing. *Trends in cognitive sciences*, 6 (2):78–84, 2002.

C. Frith and U. Frith. Theory of mind. *Current biology*, 15(17):R644–R645, 2005.

J. B. Fritz, M. Elhilali, S. V. David, and S. A. Shamma. Auditory attention—focusing the searchlight on sound. *Current opinion in neurobiology*, 17(4):437–455, 2007.

M. F. Garrett. Processes in language production. In F. J. Newmeyer, editor, *Language: Psychological and biological aspects*, pages 69–96. Cambridge University Press, 1988.

K. R. Gegenfurtner. Cortical mechanisms of colour vision. *Nature Reviews Neuroscience*, 4(7):563–572, 2003.

D. Gentner and F. Maravilla. *Analogical reasoning*, pages 186–203. The Routledge international handbook series. Routledge/Taylor & Francis Group, New York, NY, US, 2018.

S. J. Gershman and Y. Niv. Learning latent structure: carving nature at its joints. *Current opinion in neurobiology*, 20(2):251–256, 2010.

A. Gilchrist, C. Kossyfidis, F. Bonato, T. Agostini, J. Cataliotti, X. Li, B. Spehar, V. Annan, and E. Economou. An anchoring theory of lightness perception. *Psychological review*, 106(4):795, 1999.

C. Gilmore, S. M. Göbel, and M. Inglis. *An introduction to mathematical cognition*. Routledge, 2018.

S. Ginsburg and E. Jablonka. The evolution of associative learning: A factor in the cambrian explosion. *Journal of theoretical biology*, 266(1):11–20, 2010.

B. Goertzel and C. Pennachin, editors. *Artificial General Intelligence*. Cognitive Technologies. Springer Berlin, Heidelberg, 2007. ISBN 978-3-540-23733-4. doi: 10.1007/978-3-540-68677-4.

N. D. Goodman, J. B. Tenenbaum, J. Feldman, and T. L. Griffiths. A rational analysis of rule-based concept learning. *Cognitive science*, 32(1):108–154, 2008.

J. A. Grahn. Neural mechanisms of rhythm perception: current findings and future perspectives. *Topics in cognitive science*, 4(4):585–606, 2012.

A. M. Grant. Rethinking psychological mindedness: Metacognition, self-reflection, and insight. *Behaviour Change*, 18(1):8–17, 2001.

M. Gubrud. Nanotechnology and international security. In *Fifth Foresight Conference on Molecular Nanotechnology*, Nov. 1997. Presented November 1997.

L. Haas, G. Yona, G. D'Antonio, S. Goldshtein, and D. Das. Simpleqa verified: A reliable factuality benchmark to measure parametric knowledge. *arXiv preprint arXiv:2509.07968*, 2025.

R. N. Haber and M. Hershenson. *The psychology of visual perception*. Holt, Rinehart & Winston, 1973.

J. Harris and J. Smith. *Sensation and perception*. SAGE Publications Ltd, 2022.

N. Harvey. Confidence in judgment. *Trends in cognitive sciences*, 1(2):78–82, 1997.

E. Heit. Properties of inductive reasoning. *Psychonomic bulletin & review*, 7(4):569–592, 2000.

D. Hendrycks, D. Song, C. Szegedy, H. Lee, Y. Gal, E. Brynjolfsson, S. Li, A. Zou, L. Levine, B. Han, J. Fu, Z. Liu, J. Shin, K. Lee, M. Mazeika, L. Phan, G. Ingebretsen, A. Khoja, C. Xie, O. Salaudeen, M. Hein, K. Zhao, A. Pan, D. Duvenaud, B. Li, S. Omohundro, G. Alfour, M. Tegmark, K. McGrew, G. Marcus, J. Tallinn, E. Schmidt, and Y. Bengio. A definition of agi. *pre-print*, 2025.

E. T. Higgins and J. A. Bargh. Social cognition and social perception. *Annual review of psychology*, 38 (1):369–425, 1987.

A. Hollingworth. Object-position binding in visual memory for natural scenes and object arrays. *Journal of Experimental Psychology: Human Perception and Performance*, 33(1):31, 2007.

K. J. Holyoak and B. A. Spellman. Thinking. *Annual review of psychology*, 44(1):265–315, 1993.

A. Jacovi, A. Caciularu, O. Goldman, and Y. Goldberg. Stop uploading test data in plain text: Practical strategies for mitigating data contamination by evaluation benchmarks. *arXiv preprint arXiv:2305.10160*, 2023.

M. K. Johnson, S. Hashtroudi, and D. S. Lindsay. Source monitoring. *Psychological bulletin*, 114(1):3, 1993.

P. N. Johnson-Laird. Deductive reasoning. *Annual review of psychology*, 50(1):109–135, 1999.

J.-H. Jung, P. J. Seo, E. Oh, and J. Kim. Temperature perception by plants. *Trends in Plant Science*, 28 (8):924–940, 2023.

F. Katsuki and C. Constantinidis. Bottom-up and top-down attention: different processes and overlapping neural systems. *The Neuroscientist*, 20(5):509–521, 2014.

G. Kaufmann. What to measure? a new look at the concept of creativity. *Scandinavian Journal of Educational Research*, 47(3):235–251, 2003. doi: 10.1080/00313830308604. URL https://doi.org/10.1080/00313830308604.

P. Kendeou, K. L. McMaster, and T. J. Christ. Reading comprehension: Core components and processes. *Policy Insights from the Behavioral and Brain Sciences*, 3(1):62–69, 2016.

P. Kent. Fluid intelligence: A brief history. *Applied Neuropsychology: Child*, 6(3):193–203, 2017.

D. Klahr. *Exploring science: The cognition and development of discovery processes*. MIT press, 2000.

K. Kovacs and A. R. Conway. Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry*, 27(3):151–177, 2016.

A. A. Kumar. Semantic memory: A review of methods, models, and current challenges. *Psychonomic bulletin & review*, 28(1):40–80, 2021.

J. P. Leighton and R. J. Sternberg, editors. *The Nature of Reasoning*. Cambridge University Press, 2003.

A. M. Leslie, O. Friedman, and T. P. German. Core mechanisms in 'theory of mind'. *Trends in cognitive sciences*, 8(12):528–533, 2004.

D. Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Henry Holt and Co., Inc., 1982.

D. Marr and E. Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 207(1167):187–217, 1980.

F. Martínez-Plumed, R. B. Prudêncio, A. Martínez-Usó, and J. Hernández-Orallo. Item response theory in ai: Analysing machine learning classifiers at the instance level. *Artificial Intelligence*, 271:18–42, 2019. ISSN 0004-3702. doi: https://doi.org/10.1016/j.artint.2018.09.004. URL https://www.sciencedirect.com/science/article/pii/S0004370219300220.

M. G. Mattar and M. Lengyel. Planning in the brain. *Neuron*, 110(6):914–934, 2022.

R. E. Mayer and M. C. Wittrock. Problem solving. In P. A. Alexander and P. H. Winne, editors, *Handbook of Educational Psychology*, pages 287–303. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, 2006.

M. Mazeika, A. Gatti, C. Menghini, U. M. Sehwag, S. Singhal, Y. Orlovskiy, S. Basart, M. Sharma, D. Peskoff, E. Lau, J. Lim, L. Carroll, A. Blair, V. Sivakumar, S. Basu, B. Kenstler, Y. Ma, J. Michael, X. Li, O. Ingebretsen, A. Mehta, J. Mottola, J. Teichmann, K. Yu, Z. Shaik, A. Khoja, R. Ren, J. Hausenloy, L. Phan, Y. Htet, A. Aich, T. Rabbani, V. Shah, A. Novykov, F. Binder, K. Chugunov, L. Ramirez, M. Geralnik, H. Mesura, D. Lee, E.-Y. H. Cardona, A. Diamond, S. Yue, A. Wang, B. Liu, E. Hernandez, and D. Hendrycks. Remote labor index: Measuring ai automation of remote work, 2025. URL https://arxiv.org/abs/2510.26787.

M. A. McDaniel and G. O. Einstein. *Prospective memory: An overview and synthesis of an emerging field*. Sage Publications, 2007.

J. H. McDermott. The cocktail party problem. *Current Biology*, 19(22):R1024–R1027, 2009.

J. Metcalfe. Learning from errors. *Annual review of psychology*, 68(1):465–489, 2017.

K. J. Mitchell and M. K. Johnson. Source monitoring: Attributing mental experiences. In E. Tulving and F. I. M. Craik, editors, *The Oxford handbook of memory*, pages 179–195. Oxford University Press, 2000.

A. Miyake, N. P. Friedman, M. J. Emerson, A. H. Witzki, A. Howerter, and T. D. Wager. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive psychology*, 41(1):49–100, 2000.

J. Morand-Ferron. Why learn? the adaptive value of associative learning in wild populations. *Current opinion in behavioral sciences*, 16:73–79, 2017.

M. R. Morris, J. Sohl-Dickstein, N. Fiedel, T. Warkentin, A. Dafoe, A. Faust, C. Farabet, and S. Legg. Position: Levels of agi for operationalizing progress on the path to agi. In *Forty-first International Conference on Machine Learning*, 2024.

M. R. Morris, D. Altman, H. Belfield, A. Goemans, H. Iqbal, R. Burnell, I. Gabriel, S. Albanie, and A. Dafoe. Characterizing model jaggedness supports safety and usability. *Google DeepMind Technical Report*, January 2026. URL https://cs.stanford.edu/~merrie/papers/jaggedness_preprint.pdf.

T. O. Nelson. Metamemory: A theoretical framework and new findings. In G. H. Bower, editor, *The Psychology of Learning and Motivation*, volume 26 of *Psychology of Learning and Motivation*, pages 125–173. Academic Press, 1990. doi: https://doi.org/10.1016/S0079-7421(08)60053-5. URL https://www.sciencedirect.com/science/article/pii/S0079742108600535.

N. Nersessian. The cognitive basis of model-based reasoning in science. *The Cognitive Basis of Science*, 05 2002. doi: 10.1017/CBO9780511613517.008.

K. Nishikawa, A. A. Biewener, P. Aerts, A. N. Ahn, H. J. Chiel, M. A. Daley, T. L. Daniel, R. J. Full, M. E. Hale, T. L. Hedrick, A. K. Lappin, T. R. Nichols, R. D. Quinn, R. A. Satterlie, and B. Szymik. Neuromechanics: an integrative approach for understanding motor control. *Integrative and Comparative Biology*, 47(1):16–54, 05 2007. ISSN 1540-7063. doi: 10.1093/icb/icm024. URL https://doi.org/10.1093/icb/icm024.

D. Norris. Models of visual word recognition. *Trends in cognitive sciences*, 17(10):517–524, 2013.

OpenAI. Learning to Reason with LLMs, 2024. URL https://openai.com/index/learning-to-reason-with-llms/.

A. M. Owen. Cognitive planning in humans: neuropsychological, neuroanatomical and neuropharmacological perspectives. *Progress in neurobiology*, 53(4):431–450, 1997.

A. J. Oxenham. Pitch perception. *Journal of Neuroscience*, 32(39):13335–13338, 2012.

A. J. Parker. Binocular depth perception and the cerebral cortex. *Nature Reviews Neuroscience*, 8(5): 379–391, 2007.

V. Patraucean, L. Smaira, A. Gupta, A. Recasens, L. Markeeva, D. Banarse, S. Koppula, M. Malinowski, Y. Yang, C. Doersch, et al. Perception test: A diagnostic benchmark for multimodal video models. *Advances in Neural Information Processing Systems*, 36:42748–42761, 2023.

T. Patwardhan, R. Dias, E. Proehl, G. Kim, M. Wang, O. Watkins, S. P. Fishman, M. Aljubeh, P. Thacker, L. Fauconnet, N. S. Kim, P. Chao, S. Miserendino, G. Chabot, D. Li, M. Sharman, A. Barr, A. Glaese, and J. Tworek. Gdpval: Evaluating ai model performance on real-world economically valuable tasks, 2025. URL https://arxiv.org/abs/2510.04374.

L. Phan, A. Gatti, Z. Han, N. Li, J. Hu, H. Zhang, C. B. C. Zhang, M. Shaaban, J. Ling, S. Shi, et al. Humanity's last exam. *arXiv preprint arXiv:2501.14249*, 2025.

M. J. Pickering and S. Garrod. An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36(4):329–347, 2013. doi: 10.1017/S0140525X12001495.

S. Pinker. *The Language Instinct: The New Science of Language and Mind*. William Morrow and Company, New York, 1994. ISBN 978-0-06-097651-4.

D. G. Rand and M. A. Nowak. Human cooperation. *Trends in cognitive sciences*, 17(8):413–425, 2013.

M. G. Rhodes. Judgments of learning: Methods, data, and theory. In J. Dunlosky and S. K. Tauber, editors, *The Oxford Handbook of Metamemory*, pages 65–80. Oxford University Press, 06 2016. ISBN 9780199336746. doi: 10.1093/oxfordhb/9780199336746.013.4. URL https://doi.org/10.1093/oxfordhb/9780199336746.013.4.

M. Riesenhuber and T. Poggio. Models of object recognition. *Nature neuroscience*, 3(11):1199–1204, 2000.

L. J. Rips. Reasoning. *Annual review of psychology*, 41:321–353, 1990.

D. Romero-Alvarado, F. Martínez-Plumed, L. Pacchiardi, H. Save, S. M. Pawar, B. Mehrbakhsh, P. A. M. Casares, B. Slater, P. Bova, P. Romero, Z. R. Tidler, J. Prunty, L. Sun, and J. Hernandez-Orallo. Capabilities ain't all you need: Measuring propensities in ai, 2026. URL https://arxiv.org/abs/2602.18182.

J. R. Saffran. Statistical language learning: Mechanisms and constraints. *Current directions in psychological science*, 12(4):110–114, 2003.

M. Sarter, B. Givens, and J. P. Bruno. The cognitive neuroscience of sustained attention: where top-down meets bottom-up. *Brain research reviews*, 35(2):146–160, 2001.

A. Schoenfeld. *Mathematical Problem Solving*. Academic Press, 1985. ISBN 9780126288711.

D. R. Shanks. *The psychology of associative learning*, volume 13. Cambridge University Press Cambridge, 1995.

A. Shin and K. Kaneko. Large language models lack understanding of character composition of words. *arXiv preprint arXiv:2405.11357*, 2024.

B. F. Skinner. Operant behavior. *American psychologist*, 18(8):503, 1963.

M. W. Smith, J. Sharit, and S. J. Czaja. Aging, motor control, and the performance of computer mouse tasks. *Human Factors*, 41(3):389–396, 1999. doi: 10.1518/001872099779611102. URL https://doi.org/10.1518/001872099779611102.

L. K. Son and B. L. Schwartz. The relation between metacognitive monitoring and control. In T. J. Perfect and B. L. Schwartz, editors, *Applied metacognition*, pages 15–38. Cambridge University Press, 2002.

S. A. Spence, M. D. Hunter, T. F. D. Farrow, R. D. Green, D. H. Leung, C. J. Hughes, and V. Ganesan. A cognitive neurobiological account of deception: evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 359(1451):1755–1762, 2004. doi: 10.1098/rstb.2004.1555.

L. R. Squire and E. R. Kandel. *Memory : from mind to molecules*. Scientific American Library series ; no. 69. Scientific American Library, New York, 1999. ISBN 0716750716.

R. J. Stainton, editor. *Contemporary Debates in Cognitive Science*. Wiley-Blackwell, Oxford, 2006.

R. J. Sternberg and T. I. Lubart. *The Concept of Creativity: Prospects and Paradigms*, page 3–15. Cambridge University Press, 1998.

R. J. Stevenson. An initial evaluation of the functions of human olfaction. *Chemical Senses*, 35(1): 3–20, 11 2009. ISSN 0379-864X. doi: 10.1093/chemse/bjp083. URL https://doi.org/10.1093/chemse/bjp083.

E. Styles. *The psychology of attention*. Psychology Press, 2006.

R. S. Sutton, A. G. Barto, et al. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge, 1998.

H. Tajfel. Social perception. In G. Humphrey and M. Argyle, editors, *Social Psychology Through Experiment*, chapter 1, pages 20–54. Methuen, London, 1962.

P. Tarricone. *The taxonomy of metacognition*. Psychology press, 2011.

A. Taubenfeld, Z. Gekhman, L. Nezry, O. Feldman, N. Harris, S. Reddy, R. Stella, A. Goldstein, M. Croak, Y. Matias, and A. Feder. Evaluating alignment of behavioral dispositions in llms, 2026. URL https://arxiv.org/abs/2602.11328.

J. B. Tenenbaum, C. Kemp, T. L. Griffiths, and N. D. Goodman. How to grow a mind: Statistics, structure, and abstraction. *science*, 331(6022):1279–1285, 2011.

J. Theeuwes. Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & psychophysics*, 49(1):83–90, 1991.

J. T. Todd. The visual perception of 3d shape. *Trends in cognitive sciences*, 8(3):115–121, 2004.

M. J. Traxler and M. A. Gernsbacher, editors. *Handbook of Psycholinguistics*. Academic Press, London, 2 edition, 2006. ISBN 978-0-12-369374-7. doi: 10.1016/B978-0-12-369374-7.X5000-7.

A. M. Treisman. Strategies and models of selective attention. *Psychological review*, 76(3):282, 1969.

L. M. Trick and Z. W. Pylyshyn. Why are small and large numbers enumerated differently? a limited-capacity preattentive stage in vision. *Psychological Review*, 101(1):80–102, 1994. doi: 10.1037/0033-295X.101.1.80.

E. Tulving et al. Episodic and semantic memory. In *Organization of memory*, pages 381–403. Academic Press, 1972.

S. Ullman. *High-Level Vision: Object Recognition and Visual Cognition*. The MIT Press, 07 1996. ISBN 9780262285353. doi: 10.7551/mitpress/3496.001.0001. URL https://doi.org/10.7551/mitpress/3496.001.0001.

D. Van Moorselaar and H. A. Slagter. Inhibition in selective attention. *Annals of the New York Academy of Sciences*, 1464(1):204–221, 2020.

S. Vazire and E. N. Carlson. Self-knowledge of personality: Do people know themselves? *Social and personality psychology compass*, 4(8):605–620, 2010.

V. v. Veen and C. S. Carter. Conflict and cognitive control in the brain. *Current Directions in Psychological Science*, 15(5):237–240, 2006.

J. X. Wang. Meta-learning in natural and artificial intelligence. *Current Opinion in Behavioral Sciences*, 38:90–95, 2021.

R. M. Warren. *Auditory Perception*, volume 109 of *Pergamon General Psychology Series*. Pergamon, 1982. ISBN 978-0-08-025957-4.

D. B. Willingham, M. J. Nissen, and P. Bullemer. On the development of procedural knowledge. *Journal of experimental psychology: learning, memory, and cognition*, 15(6):1047, 1989.

W. Wood. Attitude change: Persuasion and social influence. *Annual review of psychology*, 51(1): 539–570, 2000.

E. Yee, E. G. Chrysikou, and S. L. Thompson-Schill. Semantic memory. In *The Oxford Handbook of Cognitive Neuroscience, Volume 1: Core Topics*. Oxford University Press, 12 2013. ISBN 9780199988693. doi: 10.1093/oxfordhb/9780199988693.013.0017. URL https://doi.org/10.1093/oxfordhb/9780199988693.013.0017.

B. T. Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology*, 2011.

N. Yeung and C. Summerfield. Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594):1310–1321, 2012.

J. M. Zacks, N. K. Speer, K. M. Swallow, and C. J. Maley. The brain's cutting-room floor: Segmentation of narrative cinema. *Frontiers in human neuroscience*, 4:168, 2010.

S. Zmigrod and B. Hommel. Feature integration across multimodal perception and action: a review. *Multisensory research*, 26(1-2):143–157, 2013.