Epistemic verbs produce spatial models

Sangeet Khemlani
sangeet.khemlani@nrl.navy.mil
Navy Center for Applied Research in Artificial Intelligence
US Naval Research Laboratory, Washington, DC 20375 USA

Abstract

Verbs such as 'know' and 'think' help people describe mental states, and reasoners without any training in logic can make epistemic inferences about mental states. For instance, verbs such as 'know' are factive, i.e., they describe true propositions, and the statement *Ora knows that it's sunny* licenses the inference that it's sunny. Logicians have accordingly developed epistemic logics capable of characterizing valid and invalid epistemic inferences based on operators that serve as analogs to verbs such as 'know' and 'think'. Recent work suggests that no existing logical system can capture the inferences that naïve individuals tend to make. This paper describes a new theory of epistemic reasoning that operates on the assumption that reasoners represent epistemic relations as spatial models. The theory accords with recent theoretical advances, existing data, as well as two novel experiments that show how reasoners cope with nested epistemic verbs, e.g., *Ami knows that Ora thinks it's sunny.*

Keywords: epistemic reasoning, mental states, knowledge, belief, factive presupposition

Introduction

"He knows nothing; and he thinks he knows everything," said Mr. Undershaft in George Barnard Shaw's play *Major Barbara*, before quipping: "That points clearly to a political career" (Shaw, 2000). People can keep track of the beliefs and knowledge of others, as Mr. Undershaft's critique shows, and scholars since antiquity have argued that doing so is a prerequisite for effective communication (Boh, 1993). Certain verbs such as *think*, *know*, *believe*, *remember, discover*, and *conclude* describe the mental states of agents, and young children are able to understand and produce discourse that includes them (e.g., Adrián, Clemente, & Villanueva, 2007; Abbeduto & Rosenberg, 1985; Booth & Hall, 1995; Forrester, 2017; Lewis, Hacquard, & Lidz, 2017; MacWhinney, 2000; Moore, 2013; Perner & Roessler, 2012; Schwanenflugel, Henderson, & Frabricius, 1998). Various literatures refer to such verbs as "propositional attitude verbs", "cognitive verbs", "mental state verbs", "epistemic verbs", "verbs of knowledge", and so on. A subset of them, i.e., verbs such as *know*, *realize*, and *discover*, are factive in nature: their objects presuppose some true condition about the world (e.g., Cohen, 1992, p. 91; Dudley, Rowe, Hacquard, & Lidz, 2017; Kiparsky & Kiparsky, 1970; Stalnaker, 1999, p. 55; but cf. Hazlett, 2012). Because factives concern truths, their presuppositions yield immediate inferences. Consider this statement:

1. Ora knows that it's sunny.

It's valid to infer that it's sunny, since it is true in any situation in which (1) is true (Jeffrey, 1981).

Epistemic verbs are often paired with complementizers, e.g., *that*, to embed independent clauses, e.g., *it's sunny*. So, epistemic verbs can be embedded recursively, as in (2):

2a. Ami knows that Ora knows that it's sunny.
b. Ami knows that Ora believes that it's cold.

These recursive embeddings can have complex and subtle inferential properties. For instance, the sentences in (2) both describe Ami's understanding about another individual's mental states. And neither licenses the inference that *Ora knows that Ami knows that it's sunny*, since epistemic relations are not symmetric (see Goodwin & Johnson-Laird, 2005). But they differ in what they presuppose: (2a) presupposes that it's sunny, but (2b) does not presuppose that it's cold, i.e., it's possible that Ora's belief is mistaken. Verbs such as *think, believe*, and *assume* are not factive, which in part explains why (2a) and (2b) license different inferences.

The difference between knowledge and belief undergirds the epistemic logics developed in the mid-20th century (Hintikka, 1962); these systems of logic stipulate the sorts of valid conclusion that follow from sentences that describe epistemic relations. Simple epistemic logics are based on the semantics of logical operators for the knowledge and belief of a particular agent, e.g., (2) above would be written as follows in orthodox epistemic logics:

3a. $K_{Ami}\ K_{Ora}\ sunny$
b. $K_{Ami}\ B_{Ora}\ cold$

where $K_{Ora}\ sunny$ symbolizes that Ora knows it's sunny, $B_{Ora}\ cold$ formalizes that Ora believes it's cold, and so on. Like many other families of formal logic, epistemic logics are based on interrelated sets of axioms. For instance, one common axiom, known as the "knowledge axiom", formalizes the notion of factivity by stipulating that the proposition P is true whenever $K_i\ P$ is true, or in other words, if an *agent i knows P*, then P is a fact. Another axiom common to orthodox epistemic logics formalizes the idea of "logical omniscience", which states that if $K_i\ P$ is true, and if P implies Q, then $K_i\ Q$ is true too, i.e., that agents know every implication of their knowledge. The axiom is not psychologically plausible, as logicians recognize (Hintikka, 1962, p. 30), but systems of formal logic can be useful in the absence of psychological plausibility (Fagin, Halpern, Moses, & Vardi, 2003). More complex systems of epistemic logic (e.g., Baltag & Renne, 2016; Hendricks, 2006; Meyer & van Der Hoek, 2004; Van Ditmarsch, van Der Hoek, & Kooi, 2007) characterize how states of knowledge change between multiple agents, though most of them operate on a common set of axioms (i.e., the "S5" axioms; see Rendsvig & Symons, 2019). The systems were intended to explore ideal reasoning, but because they rely on certain implausible assumptions (such as logical omniscience), they do not serve as a feasible account of human mental state reasoning (Ragni & Johnson-Laird, 2018, 2019).

Psychologists have conducted extensive programs of research into aspects of epistemic reasoning, such as the ability to understand false beliefs (e.g., Gopnik & Wellman, 1992; Perner, Huemer, & Leahy, 2015; Saxe, Carey, & Kanwisher, 2004) and certain egocentric biases that occur when individuals know things that others don't know (Birch & Bloom, 2007; Surtees & Apperly, 2012). And research into the acquisition of verbs such as *know*, *think*, *guess*, and *believe* (e.g., Kuhn, 1989; Kuhn et al., 1995; Montgomery, 1992; Sodian & Wimmer, 1987) shows that epistemic verbs mature at different rates, e.g., children understand verb *know* before they understand *think*, and the verb *guess* matures well into later childhood (Abbeduto & Rosenburg, 1985). Recent theoretical frameworks posit the centrality of knowing over believing (Philipps et al., 2020).

Despite these discoveries, no theory explains the mental representations and processes that underlie the human reasoning about others' mental states. In particular, accounts of theory of mind in children and adults concern the relative difficulty of false beliefs, i.e., when an individual knows that another agent's belief is wrong. They don't predict or explain more prosaic patterns of epistemic reasoning, such as how people mentally represent or reason about the statements in (2) above. For that reason, experiments have not investigated any relative difficulty in how people process such statements.

The present paper addresses the deficit. It first spells out a new theory of the mental representations that underlie the verbs *know* and *think*: it argues that people construct small-scale spatial simulations to represent others' mental states. It shows how people can perform directional scans on the models to make epistemic inferences about who knows what. It then derives two predictions from the theory, and reports on experiments that tested and corroborated them. The theory and the experiments open new avenues for research into mental state reasoning, and the paper concludes with a discussion of existing controversies and how future research can address them.

Spatial models of *know* and *think*

A viable theory of epistemic reasoning must explain two central patterns of mental state inference: the first is factivity, i.e., that some epistemic verbs help reasoners make inferences about the real world, not just mental states. The second is that people can have knowledge about knowledge, i.e., knowledge about whether another individual has knowledge, as in (2a). At first blush, the second phenomenon may seem to suggest that representations of mental states are recursive in nature (see Figure 1 for an illustration). But recursive structures pose problems for embedded epistemic operators, as in (2a). If (2a) is recursively embedded, as in Figure 1a, then accessing the presupposed fact – that it's sunny – requires individuals to first access Ami's knowledge

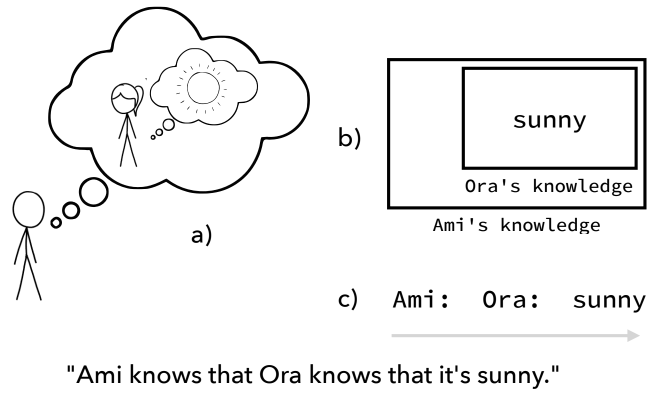

Figure 1. How might reasoners represent the statement in (2a)? The statement concerns the mental state of an agent, Ami, depicted graphically in (a). It can be represented as a recursive structure, as in (b). It can also be represented spatially along a single dimension (c), i.e., a spatial model. The latter representation can be scanned directionally to yield epistemic inferences.

about Ora, then Ora's knowledge about the state of the world. The complexity of the procedure seems at odds with the simplicity of the inference: reasoners should not find it difficult to infer that it's sunny from statements such as (2a). Likewise, recursive structures predict an increase in difficulty with each additional recursive relation, i.e., it should be easier to make inferences from (2a) than, e.g., the statement: *Dev knows that Ami knows that Ora knows that it's sunny*. Finally, it is not clear how recursive representations support simple epistemic inferences, such as the inference that *Ami knows that it's sunny*.

An alternative account is that reasoners make epistemic inferences by building mental simulations of epistemic relations. A prominent theory of mental simulation argues that people build simulations of possibilities – mental models – that mimic the structure of the situations they represent (Johnson-Laird, 2006, 2010). Models explain how people reason about time (Schaeken et al., 1996; Kelly, Khemlani, & Johnson-Laird, 2020), causality (Goldvarg & Johnson-Laird, 2001; Johnson-Laird & Khemlani, 2017; Khemlani et al., 2021), kinematics (Khemlani, Mackiewicz, Bucciarelli, & Johnson-Laird, 2013), and various other kinds of relation (Goodwin & Johnson-Laird, 2005). For instance, Ragni and Knauff (2013; see also Knauff, 2013) show how reasoners construct models to compute spatial inferences. To interpret, e.g., descriptions such as:

4. The farmer is to the left of the baker.
   The brewer is to the right of the farmer.

reasoners construct small scale spatial models of the relevant scenario, e.g.,

```
farmer    brewer    baker
```

Such models represent what is common to any real-world situation consistent with (4). And they can be scanned to yield inferences, e.g., scanning the model above from left to right yields the simple deduction that *the brewer is to the right of the baker.*

Spatial models – indeed, models in general – help explain many patterns of human reasoning, such as why reasoners make systematic errors (e.g., Ragni, Sonntag, & Johnson-Laird, 2016) and why they exhibit predictable biases (Jahn, Knauff, & Johnson-Laird, 2007). Moreover, models help explain how reasoners without any training in formal logic and other cognate disciplines can make valid inferences.

The processes that underlie spatial reasoning may apply to epistemic reasoning as well. I posit that reasoners represent statements such as (1) above as spatial models that concern an agent and a particular state of knowledge, e.g.,

```
Ora:  sunny
```

The first token – `Ora:` – denotes that anything that follows is a part of Ora's knowledge. The second token, `sunny`, represents what Ora knows. Hence, a spatial model of (2a) is:

```
Ami:   Ora:  sunny
```

Unlike typical spatial models, epistemic models can only be scanned in a single direction, i.e., from a knower (e.g., Ami) to the knower's knowledge. Hence, a directional scan yields the following acceptable inferences:

Ora knows it's sunny.
Ami knows it's sunny.
It's sunny.

But it does not yield these inferences:

Ora knows that Ami knows it's sunny.
Ami knows that Ora knows that Ami knows it's sunny.

and so on, which is just as well, since none of them are valid.

How do reasoners separate between factive verbs, such as *know*, and non-factive verbs, such as *think*? The theory argues that they represent a token that *suppresses* a directional scan. For example, reasoners should represent the statement, *Ami knows that Ora thinks that it's sunny*, as follows:

```
Ami:   Ora: *sunny
```

The model is identical to the model reasoners represent for (2a), except that a token suppresses the inference that it's sunny. Hence, the model permits only this inference, which is valid:

Ora thinks it's sunny.

And it does not permit any of the following invalid inferences:

It's sunny.
Ami knows that it's sunny.
Ami thinks that it's sunny.
Ora thinks that Ami thinks that it's sunny.

One consequence of the theory is that it can make predictions about unexplored patterns of epistemic reasoning. Consider the mental model people might build for the statement, *Ami thinks that Ora knows that it's sunny*, which is as follows:

```
Ami:   *Ora:  sunny
```

The statement embeds a factive proposition, i.e., "Ora knows that it's sunny", so orthodox linguistic analyses (e.g., Kiparsky & Kiparsky, 1970) might suggest that the statement presupposes that *it's sunny*. But, if reasoners make inferences based on epistemic models, and if they scan such models from *knower* (i.e., Ami) to *knowledge* (i.e., that it's sunny) to draw inferences, some reasoners may mistakenly process the suppression token as they scan the model. And doing so may cause them to suppress the inference that it's sunny. Hence, the epistemic model theory makes the following prediction:

Prediction: People should infer P from *X knows that Y knows that P* more often than for *X thinks that Y knows that P*.

In what follows, I describe two experiments that tested and validated the prediction.

Experiment 1

Experiment 1 tested the prediction above. It presented participants with problems such as the following:

Suppose that Jesse and Winter come across a particular animal. Jesse thinks that Winter knows that the animal is a toad.

Do you agree with the following statement?
The animal is a toad.

The study varied the second premise of the problems, which were in one of four sentence forms:

X knows that Y knows that P.	[knows/knows]
X knows that Y thinks that P.	[knows/thinks]
X thinks that Y knows that P.	[thinks/knows]
X thinks that Y thinks that P.	[thinks/thinks]

where *X* and *Y* stand in place of the names of individuals, e.g., "Jesse" and "Winter", and where *P* stands in place of a sentence describing an animal, e.g., "The animal is a toad." Likewise, it varied the conclusion that participants had to evaluate, which could have been in one of three sentence forms:

P.	[presupposition]
X knows that P.	[immediate inference]
Y knows that P.	[immediate inference]

The study therefore generated 12 separate problems, and each participant carried out all 12. Accounts based on factivity alone should predict no difference between *knows/knows* and *thinks/knows* problems, and no difference between *knows/thinks* and *thinks/thinks* problems, because *P* in each case should depend on only the factivity of the epistemic verb. The model theory predicts otherwise.

Method

Participants. 52 participants completed the experiment for monetary compensation ($2 and a 10% chance of a $10 bonus) through Amazon Mechanical Turk. All of the participants stated that they were native English speakers.

Open science. Data, materials, experimental code, and analysis scripts for both this experiment and the following one are available through the Open Science Framework platform (https://osf.io/q57tk/).

Design and materials. Participants acted as their own controls and assessed whether a conclusion holds for 12 separate problems. The experiment manipulated the epistemic verbs in statements such as:

> Kamryn [knows/thinks] that London [knows/thinks] that the animal is a katydid.

to yield four separate types of problems: *knows/knows*, *knows/thinks*, *thinks/knows*, and *thinks/thinks*. For each problem, the names of the individuals, e.g., "Kamryn" and "London" were randomized, unique, and non-gendered. Likewise, the animals in the study were unique for each problem, and randomized such that no participant ever saw the same combination of names of individuals with animals. Participants pressed a button marked "yes" or "no" to indicate whether they agreed with a given conclusion. The conclusions were in one of three forms:

> The animal is a katydid.
> Kamryn knows that the animal is a katydid.
> London knows that the animal is a katydid.

Hence, the study yielded a 4 (problem types) x 3 (conclusion forms) fully repeated-measures design. The 12 resulting problems occurred in a different random order for each participant.

Results and discussion

Table 1 presents the percentages of participants' "yes" responses as a function of the four types of problems and the three types of conclusions in the experiment. Friedman nonparametric analyses of variance revealed that their responses differed as a function of the four different problem types (Friedman test, χ^2 = 81.39, p < .0001), but not as a function of the different conclusion types (Friedman test, χ^2 = 3.62, p = .16). The results suggest that reasoners who infer a particular conclusion, e.g., *the animal is a katydid*, are likely to also infer that *Kamryn knows that the animal is a katydid* and *London knows that the animal is a katydid*. Subsequent Pearson correlations revealed that agreement data between the three types of conclusion were highly intercorrelated (Pearson tests, rs > .45, ps < .0001). Hence, data for the separate conclusions were pooled for subsequent planned comparisons.

An initial planned comparison confirmed that participants responded sensibly throughout the experiment: they accepted conclusions more often for the three problems that contained one or more instances of the verb 'know' than for the problem that only contained the verb 'think' (65% vs. 28%, Wilcoxon test, z = 5.69, p < .0001, Cliff's δ = 0.62). Additional planned comparisons tested the predictions of the spatial model theory. Participants accepted reliably more conclusions for *knows/knows* problems than *thinks/knows* problems (95% vs. 62%, Wilcoxon test, z = 5.51, p < .0001, Cliff's δ = 0.57), which corroborates the prediction outlined above. And they accepted more conclusions for *knows/thinks* compared to *thinks/thinks* (36% vs. 28%, Wilcoxon test, z = 2.34, p = .019, Cliff's δ = 0.10). Hence, Experiment 1 validated the prediction of the spatial model theory.

Table 1. The percentages of participants' 'yes' responses to conclusions as a function of the four types of problems and the three types of conclusions they evaluated across Experiment 1.

	Conclusion type		
Problem type	P	X knows that P	Y knows that P
X knows that Y knows that P	96	98	90
X knows that Y thinks that P	37	38	35
X thinks that Y knows that P	71	58	60
X thinks that Y thinks that P	35	25	25

In Experiment 1, participants provided binary responses, i.e., "yes" or "no", to assess each conclusion, and those responses may have ridden roughshod over their intuitions. For instance, the task may have inflated agreement to problems of the form *X thinks that Y thinks that P,* because reasoners who do may think *P* might be possible were forced to agree with the statement in its entirety. Hence, Experiment 2 replicated the design of Experiment 1, but it allowed participants to register their responses on a 7-point Likert scale.

Experiment 2

Experiment 2 was similar to Experiment 1 in all respects, except that participants responded to a different question, namely: "To what extent do you agree with the following statement"? They registered their responses by manipulating a slider to mark their preference on a 7-point Likert scale. Hence, the experiment was a 4 x 3 fully repeated measures design.

Method

Participants. 50 participants completed the experiment on Amazon Mechanical Turk for monetary compensation ($2 and a 10% chance of a $10 bonus). All of the participants stated that they were native English speakers.

Open science. The OSF link is stated in Experiment 1.

Design, materials, and task. The design and materials of the Experiment 2 were the same as in Experiment 1. Instead of a binary response option, participants registered their responses using a 7-point Likert scale (-3 = *definitely false*, 0 = *I cannot be certain*, +3 = *definitely true*). At the start of each problem, the scale's value defaulted to 0. Participants were not permitted to move on to the next problem before clicking on the scale to mark their choice.

Results and discussion

Table 2 presents the percentages of participants' mean ratings as a function of the four types of problems and the three types of conclusions in the experiment. The study yielded a significant main effect of the type of problem (Friedman test, χ^2 = 81.39, p < .0001), but and a marginal effect of the different conclusion types (Friedman test, χ^2 = 4.78, p = .09). The results replicate the previous study, and further establish that reasoners tend to rate inferences such as *the animal is a katydid* as similarly acceptable to inferences such as *Kamryn knows that the animal is a katydid* and *London knows that the animal is a katydid*. Pearson correlations revealed that highly correlated ratings between the three types of conclusion (Pearson tests, rs > .39, ps < .004). As in Experiment 1, data for the separate conclusions were pooled for subsequent planned comparisons.

A planned comparison confirmed that participants understood the task and responded in a manner that reflected intuitions about factives, i.e., they rated conclusions higher for the three problems that contained one or more instances of the verb 'know' than for the problem that only contained the verb 'think' (.95 vs. -0.04, Wilcoxon test, z = 3.39, p < .0001, Cliff's δ = 0.41). Participants rated conclusions to *knows/knows* problems higher than *thinks/knows* problems (1.91 vs. .57, Wilcoxon test, z = 5.09, p < .0001, Cliff's δ = 0.64) and thereby corroborated the theory's prediction. They also rated conclusions for *knows/thinks* higher than for *thinks/thinks*, also in line with the model theory processing constraints, but the pattern was not reliable (.00 vs. -.27, Wilcoxon test, z = 0.80, p = .42, Cliff's δ = 0.12).

In sum, Experiments 1 and 2 corroborated the model theory of epistemic reasoning, which shows how mental simulations can be scanned to make mental state inferences.

General discussion

A viable theory of epistemic reasoning must explain at least two things: how people distinguish factives from non-factives, and how they mentally represent embedded epistemic relations. Indeed, many researchers have examined specific scenarios that concern embedded epistemic relations: false belief tasks concern a scenario in which one agent knows that another agent's belief is false (Gopnik & Wellman, 1992; Perner, Huemer, & Leahy, 2015; Saxe, Carey, & Kanwisher, 2004). But no general theory explains how adults reason about epistemic relations such as *know* and *think*. The theory I describe reduces epistemic reasoning to a form of spatial inference: it posits that reasoners construct spatial models of epistemic relations, and that they scan those models directionally to yield systematic inferences.

The advantages of the theory are at least two-fold. First, the theory explains how reasoners make common epistemic inferences, e.g., it explains how they infer that *it's sunny* from statements of the form, *Ami knows that Ora knows that it's sunny*. They build a spatial model, as depicted in this diagram:

```
Ami:   Ora:  sunny
```

and they scan the model from knower (i.e., Ami) to knowledge (i.e., that it's sunny). This procedure explains not just common inferences, but also how people avoid drawing invalid conclusions, e.g., that *Ora knows that Ami knows that it's sunny*: they do not scan the model in a way that permits such an inference.

Second, the theory explains how people make certain mistakes in epistemic reasoning. Linguists argue that this sentence: *Ami thinks that Ora knows that it's sunny*, unequivocally presupposes that it's sunny, but reasoners in two experiments did not draw such a conclusion as often as they did for the sentence *Ami knows that Ora knows that it's sunny*, which yields the same presupposition. No theory until now could predict or explain such differences in performance.

The model theory of epistemic reasoning accords with a recent theoretical framework proposed by Phillips and colleagues (2020). They marshal data from studies on non-human primates, children, and non-neurotypical populations to argue that representations of knowledge are more primitive than representations of belief. The theory presented above separates knowledge and belief (linguistically realized by the use of verbs such as *know* and *think*) based on the different epistemic models that reasoners build to process such verbs. Epistemic models of the verb *thinks* are complex. The statement, *Ami thinks that Ora knows that it's sunny* is:

```
Ami:   *Ora:  sunny
```

where the * symbol indicates a point at which directional scans are suppressed. This difference may explain, e.g., why children learn and use *know* before they use *think* (Abbeduto & Rosenburg, 1985).

I highlight one final advantage of the theory: certain neurocognitive measures can be used to trace the kinds of spatial models reasoners build when they engage in spatial reasoning tasks (Alfred, Connolly, Cetron, & Kraemer, 2020; Fangmeier, Knauff, Ruff, & Sloutsky, 2006; Knauff, 2009; Ragni, Franzmeier, Maier, & Knauff, 2016). If epistemic reasoning is a form of specialized spatial inference, then imaging data may help assess the theory I outline and its novel predictions.

Table 2. The mean of participants' belief ratings (which ranged from -3 to +3) as a function of the four types of problems and the three types of conclusions they evaluated in Experiment 2.

	Conclusion type		
Problem type	P	X knows that P	Y knows that P
X knows that Y knows that P	2.00	1.70	2.02
X knows that Y thinks that P	0.10	-0.02	-0.08
X thinks that Y knows that P	0.74	0.70	0.28
X thinks that Y thinks that P	-0.04	-0.28	-0.48

Acknowledgments

This research was supported by a grant from the Office of Naval Research to SK. I'm grateful to Kalyan Gupta, Danielle Paterno, and Kevin Zish at the Knexus Research Corporation for their help in conducting the experiments. I also thank Bill Adams, Gordon Briggs, Ruth Byrne, Hillary Harner, Tony Harrison, Laura Hiatt, Zach Horne, Phil Johnson-Laird, Laura Kelly, Joanna Korman, Markus Knauff, Andrew Lovett, Marco Ragni, Greg Trafton, and Ben Wright for their suggestions and critiques.

References

Abbeduto, L., & Rosenberg, S. (1985). Children's knowledge of the presuppositions of know and other cognitive verbs. *Journal of Child Language,* 12.

Adrián, J., Clemente, R.A, & Villanueva, L. (2007). Mothers' use of cognitive state verbs in picture-book reading and the development of children's understanding of mind: A longitudinal study. *Child Development,* 78.

Alfred, K. L., Connolly, A. C., Cetron, J. S., & Kraemer, D. J. (2020). Mental models use common neural spatial structure for spatial and abstract content. *Communications Biology*, 3.

Baltag, A. & Renne, B. (2016). Dynamic epistemic logic. In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy.* Stanford University.

Birch, S. A., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, 18.

Boh, I. (1993). *Epistemic logic in the Later Middle Ages*. New York, NY: Routledge.

Booth, J. R., & Hall, W. S. (1995). Development of the understanding of the polysemous meanings of the mental-state verb know. *Cognitive Development*, 10.

Cohen, L. J. (1992). *An essay on belief and acceptance.* Oxford University Press.

Dudley, R., Rowe, M., Hacquard, V., & Lidz, J. (2017). Discovering the factivity of *know*. *Semantics and Linguistic Theory*, 27.

Fagin, R., Moses, Y., Halpern, J. Y., & Vardi, M. Y. (2003). *Reasoning about knowledge.* Cambridge, MA: MIT Press.

Fangmeier, T., Knauff, M., Ruff, C. C., & Sloutsky, V. (2006). fMRI evidence for a three-stage model of deductive reasoning. *Journal of Cognitive Neuroscience*, 18.

Forrester, M. (2017). Learning how to use the word 'know': examples from a single-case study. In *Children's Knowledge-in-Interaction*. Springer, Singapore.

Goldvarg, E., & Johnson-Laird, P. N. (2001). Naive causality: A mental model theory of causal meaning and reasoning. *Cognitive Science*, 25.

Goodwin, G., & Johnson-Laird, P.N. (2005). Reasoning about relations. *Psychological Review,* 112.

Gopnik, A., & Wellman, H. (1992). Why the child's theory of mind really is a theory. *Mind & Language,* 7.

Hazlett, A. (2012). Factive presupposition and the truth condition on knowledge. *Acta Analytica,* 27.

Hendricks, V.F. (2006): Mainstream and Formal Epistemology, New York: Cambridge University Press.

Hintikka, J. (1962). *Knowledge and belief.* Ithaca, NY: Cornell University Press.

Jahn, G., Knauff, M., & Johnson-Laird, P. N. (2007). Preferred mental models in reasoning about spatial relations. *Memory & Cognition*, 35.

Jeffrey, R. (1981). *Formal logic: Its scope and limits (2nd Ed).* New York: McGraw-Hill.

Johnson-Laird, P. N. (2006). *How we reason.* Oxford University Press, USA.

Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proceedings of the National Academy of Sciences,* 107.

Johnson-Laird, P. N., & Khemlani, S (2017). Mental models and causation. In M. Waldmann (Ed.), *Oxford Handbook of Causal Reasoning*.

Kelly, L. J., Khemlani, S., & Johnson-Laird, P. N. (2020). Reasoning about durations. *Journal of Cognitive Neuroscience*, 32.

Khemlani, S., Mackiewicz, R., Bucciarelli, M., & Johnson-Laird, P.N. (2013). Kinematic mental simulations in abduction and deduction. *Proceedings of the National Academy of Sciences*, 110.

Kiparsky, P., & Kiparsky, C. (1970). Fact. In M. Bierwisch & K.E. Heidolph (Eds.), *Progress in linguistics: A collection of papers* (pp. 143–173). The Hague: Mouton.

Knauff, M. (2009). A neuro-cognitive theory of deductive relational reasoning with mental models and visual images. *Spatial Cognition & Computation*, 9.

Knauff, M. (2013). *Space to reason: A spatial theory of human thought.* MIT Press.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96.

Kuhn, D., Garcia-Mila, M., Zohar, A., Andersen, C., White, S. H., Klahr, D., & Carver, S. M. (1995). Strategies of knowledge acquisition. *Monographs of the Society for Research in Child Development*, i-157.

Lewis, S., Hacquard, V., & Lidz, J. (2017). "Think" pragmatically: Children's interpretation of belief reports. *Language Learning and Development,* 13.

Perner, J., Huemer, M., & Leahy, B. (2015). Mental files and belief: A cognitive theory of how children represent belief and its intensionality. *Cognition*, 145.

Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences,* 16, 519-525.

Phillips, J., et al. (2020). Knowledge before belief. *Behavioral and Brain Sciences.*

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk, transcription format and programs* (3rd ed). Mahwah, NJ: Lawrence Erlbaum Associates.

Meyer, J. J. C., & Van Der Hoek, W. (2004). *Epistemic logic for AI and computer science* (No. 41). Cambridge University Press.

Montgomery, D. (1992). *Young children's theory of knowing: The development of a folk epistemology. Developmental Review*, 12.

Moore, C. (2013). *The development of commensense psychology.* Psychology Press.

Ragni, M., & Johnson-Laird, P.N. (2018). Reasoning about possibilities: Human reasoning violates all normal modal logics. In *Proceedings of the Annual Conference of the Cognitive Science Society.*

Ragni, M., & Johnson-Laird, P.N. (2019). Possibilities as the foundation of reasoning. *Cognition,* 193.

Ragni, M., & Knauff, M. (2013). A theory and a computational model of spatial reasoning with preferred mental models. *Psychological Review*, 120.

Ragni, M., Franzmeier, I., Maier, S., & Knauff, M. (2016). Uncertain relational reasoning in the parietal cortex. *Brain & Cognition*, 104.

Ragni, M., Sonntag, T., & Johnson-Laird, P. N. (2016). Spatial conditionals and illusory inferences. *Journal of Cognitive Psychology*, 28.

Rendsvig, R. & Symons, J. (2019). *Epistemic logic.* In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Stanford University.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, 55.

Schaeken, W., Johnson-Laird, P. N., & d'Ydewalle, G. (1996). Mental models and temporal reasoning. *Cognition*, 60.

Schwanenflugel, P., Henderson, R., & Fabricius, W. (1998). Developing organization of mental verbs and theory of mind in middle childhood: Evidence from extensions. *Developmental Psychology,* 34, 512-524.

Sodian, B., & Wimmer, H. (1987). Children's understanding of inference as a source of knowledge. *Child Development*, 424-433.

Stalnaker, R. (1999). *Context and content.* Oxford University Press.

Surtees, A. D., & Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development*, 83, 452-460.

Van Ditmarsch, H., van Der Hoek, W., & Kooi, B. (2007). *Dynamic epistemic logic* (Vol. 337). Springer Science & Business Media.