LINGUISTIC COMMUNICATION & SOCIAL COGNITION

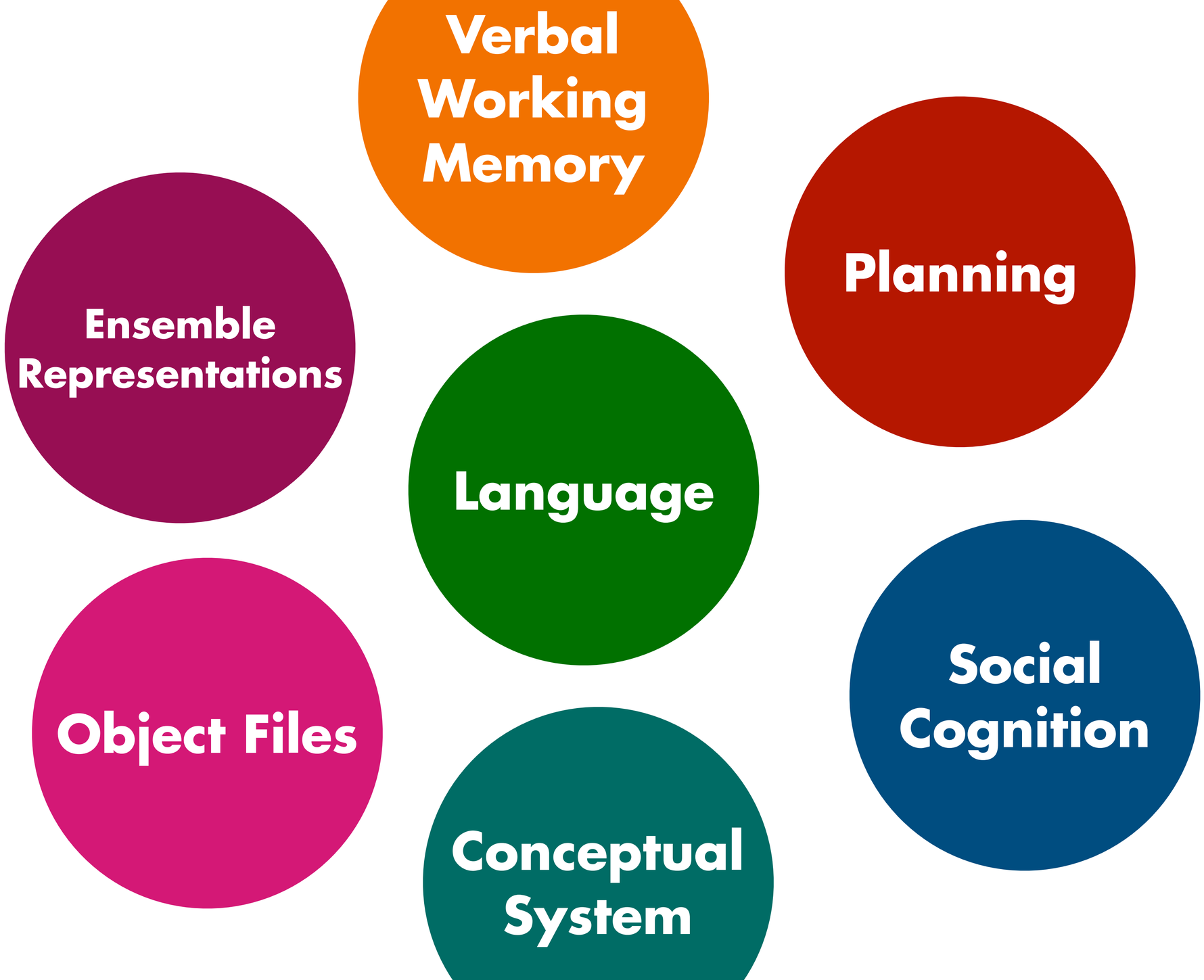
PETER VAN ELSWYK
NORTHWESTERN

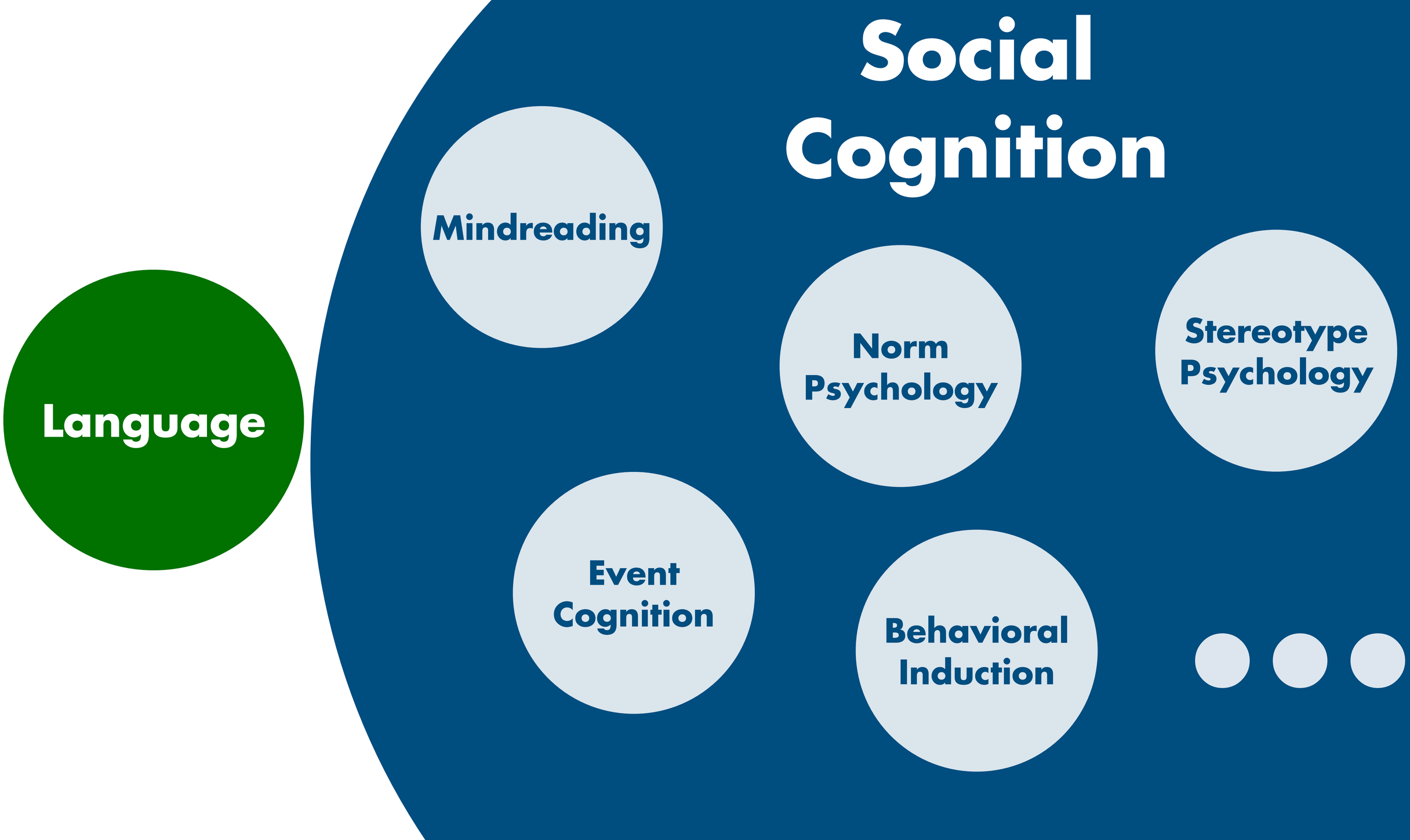
DANIEL HARRIS
CUNY GRADUATE CENTER, HUNTER COLLEGE

DAY 5:
FACTIVE MINDREADING

PETER VAN ELSWYK
NORTHWESTERN

DANIEL HARRIS
CUNY GRADUATE CENTER, HUNTER COLLEGE


Verbal Working Memory
Planning
Ensemble Representations
Language
Social Cognition
Object Files
Conceptual System


Social Cognition
Mindreading
Norm Psychology
Stereotype Psychology
Language
Event Cognition
Behavioral Induction

Mindreading

Language

Belief Attribution

Minimal Mindreading

Emotion Attribution

Knowledge Attribution

Goal Attribution

Plan Attiribution

Language

Mindreading

Belief Attribution

Minimal Mindreading

Emotion Attribution

Knowledge Attribution

Goal Attribution

Plan Attiribution

Knowledge Attribution

An attitude in which an agent non-accidentally represents the truth.

Belief Attribution

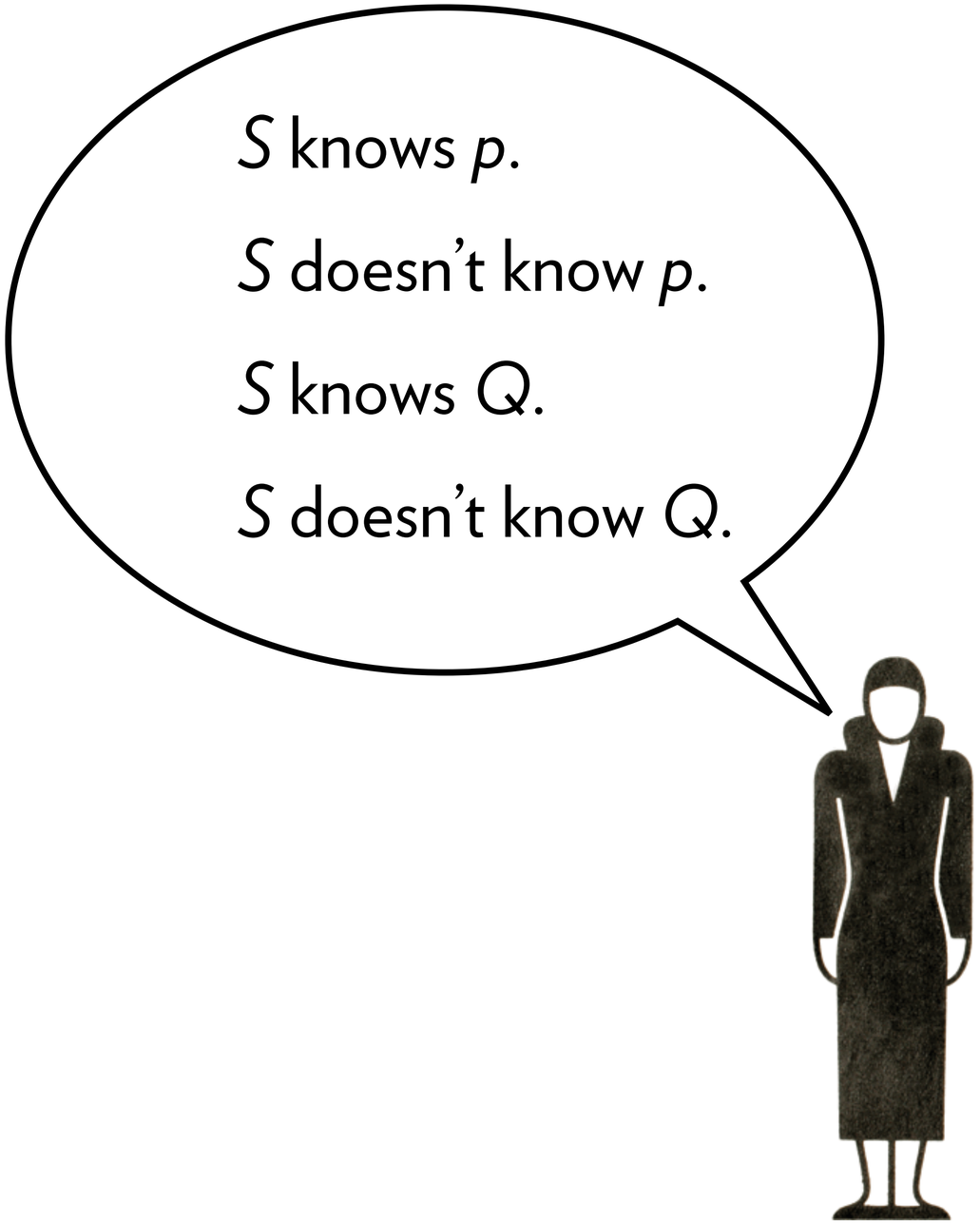
An attitude in which an agent represents.

Knowledge Attribution

An attitude in which an agent non-accidentally represents the truth.

Belief Attribution

An attitude in which an agent represents.

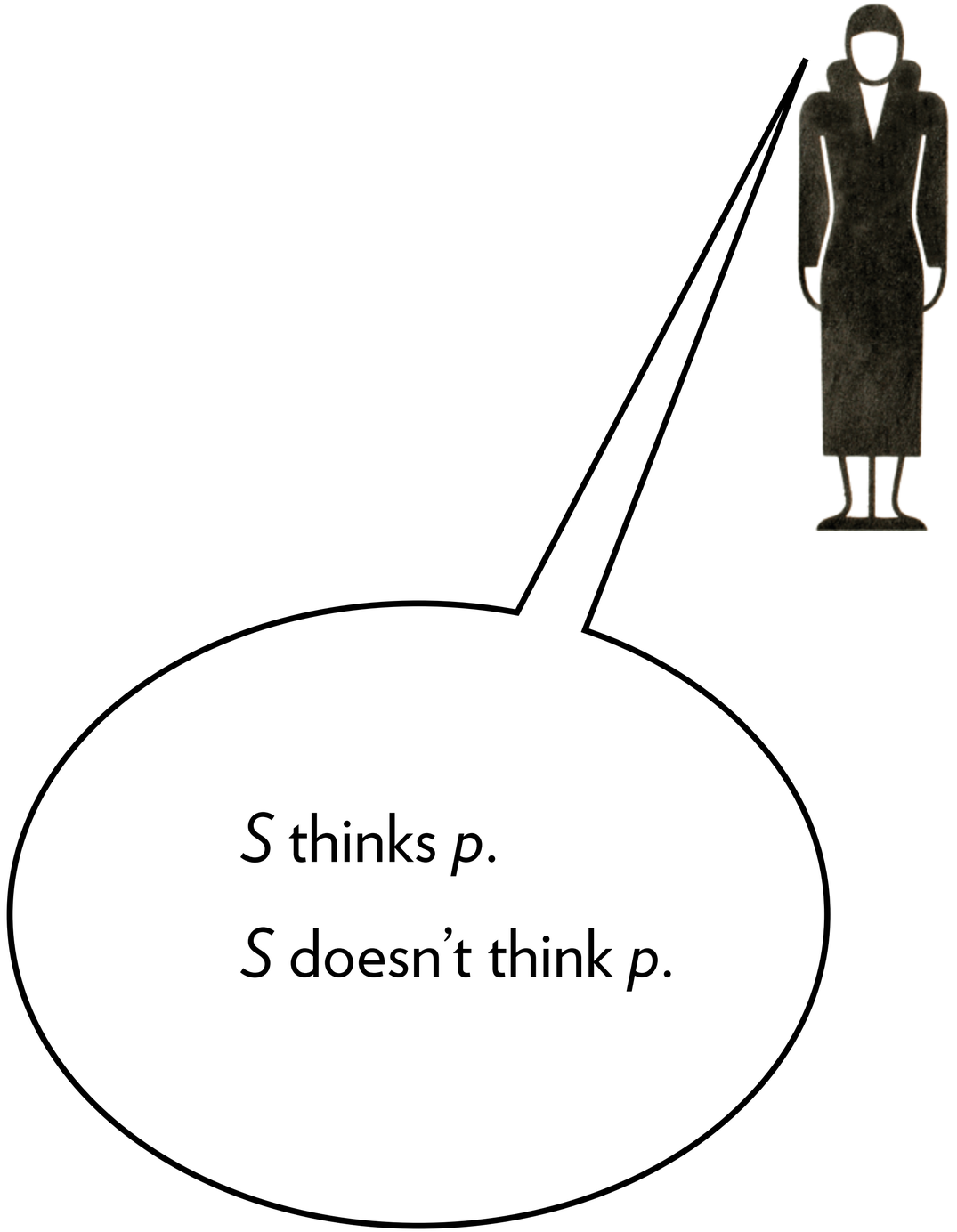


Knowledge Attribution

An attitude in which an agent non-accidentally represents the truth.

Belief Attribution

An attitude in which an agent represents.



WHAT IS THE DIFFERENCE BETWEEN ATTRIBUTING KNOWLEDGE AND BELIEF?

Confusion Warning

Confusion #1

Mistaking *which attitude is attributed* with *which attitude an agent has*.

Confusion #1

Mistaking *which attitude is attributed* with *which attitude an agent has*.

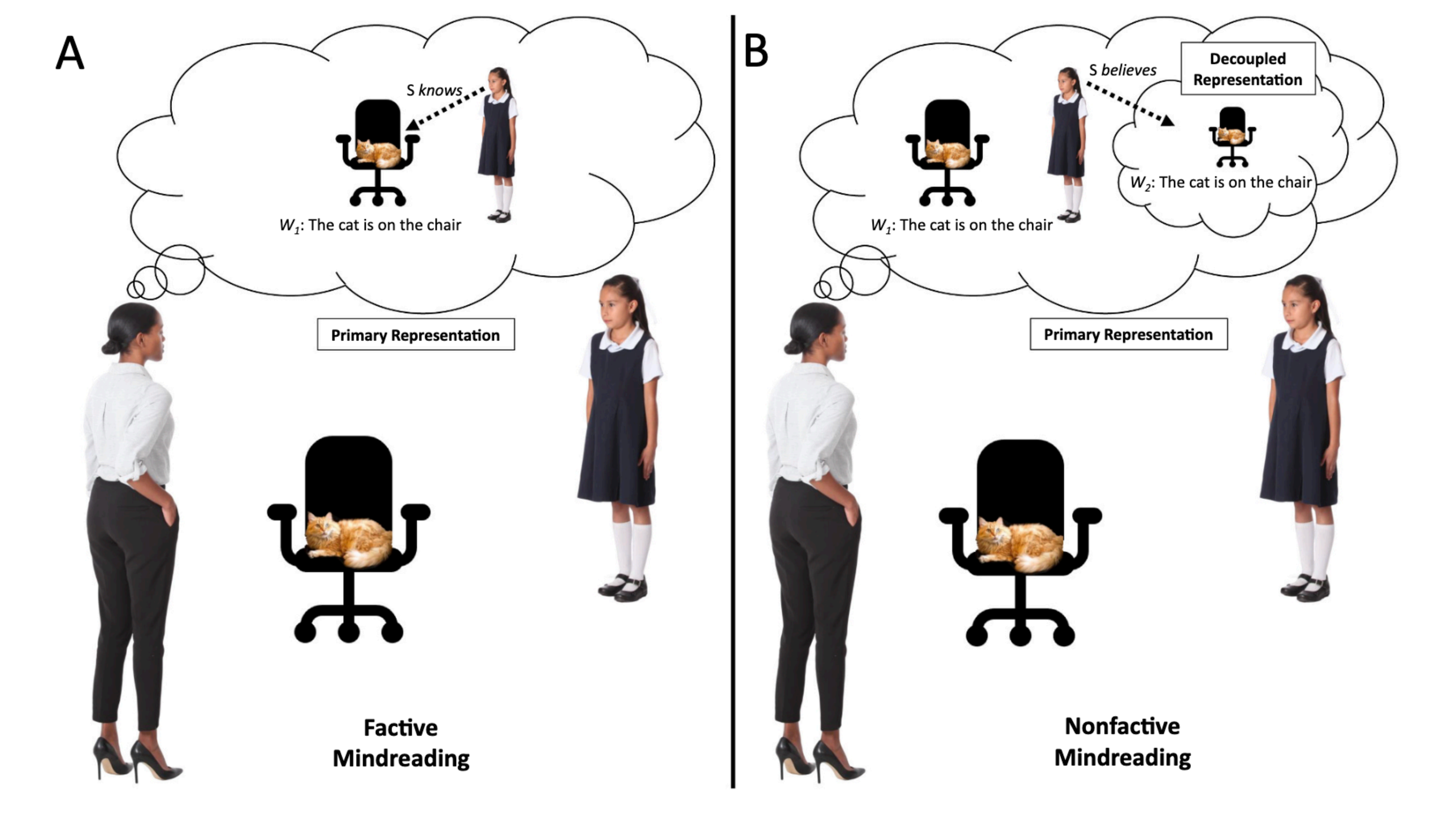
Confusion #2

Mistaking *how knowledge is represented* with *how knowledge is traditionally analyzed in philosophy*

Attitude representation

"*Factive mentalizing* is employed when we represent others as either knowing or failing to know about some aspect of reality. *Nonfactive mentalizing* governs the attribution of beliefs whose contents are decoupled from the attributor's larger representation of what is real."

— Evan Westera and Jennifer Nagel


A
S knows
W1: The cat is on the chair
Primary Representation
Factive
Mindreading
B
S believes
Decoupled
Representation
W2: The cat is on the chair
W1: The cat is on the chair
Primary Representation
Nonfactive
Mindreading

Factive theory of mind

Jonathan Phillips[1] | **Aaron Norby**[2]

[1]Department of Psychology, Harvard University, Cambridge, Massachusetts

[2]Department of Philosophy, Yale University, New Haven, Connecticut

Correspondence
Jonathan Phillips, 1482 William James Hall, 33 Kirkland Street, Cambridge, MA 02138.
Email: phillips01@g.harvard.edu

Research on theory of mind has primarily focused on demonstrating and understanding the ability to represent others' non-factive mental states, for example, others' beliefs in the false-belief task. This requirement confuses the ability to represent a particular kind of non-factive content (e.g., a false belief) with the more general capacity to represent others' understanding of the world even when it differs from one's own. We provide a way of correcting this. We first offer a simple and theoretically motivated account on which *tracking* another agent's understanding of the world and

1 | INTRO

On an island ir
approached free-
Nissen & Ferrug
on the ground an

Behavioral and Brain Sciences

cambridge.org/bbs

Knowledge before belief

Jonathan Phillips[a], Wesley Buckwalter[b], Fiery Cushman[c], Ori Friedman[d], Alia Martin[e], John Turri[f], Laurie Santos[g], and Joshua Knobe[h]

[a]Program in Cognitive Science, Department of Psychological and Brain Sciences and Department of Philosophy, Dartmouth College, Hanover, NH 03755, USA; [b]Department of Philosophy, Institute for Philosophy and Public Policy, George Mason University, Fairfax, VA 22030, USA; [c]Department of Psychology, Harvard University, Cambridge, MA 02138, USA; [d]Department of Psychology, University of Waterloo, Waterloo, ON, N2L 3G1, Canada; [e]School of Psychology, Victoria University of Wellington, Wellington, 6012, New Zealand; [f]Philosophy Department and Cognitive Science Program, University of Waterloo, Waterloo, ON N2L 3G1 Canada; [g]Department of Psychology, Yale University, New Haven, CT 06520, USA and [h]Program in Cognitive Science, Department of Philosophy, Yale University, New Haven, CT 06520, USA.
jonathan.s.phillips@dartmouth.edu; http://phillab.host.dartmouth.edu/
wesleybuckwalter@gmail.com; https://wesleybuckwalter.org/
cushman@fas.harvard.edu; http://cushmanlab.fas.harvard.edu/
friedman@uwaterloo.ca; https://sites.google.com/view/uwaterloocclab
alia.martin@vuw.ac.nz; https://vuwbabylab.com/
john.turri@gmail.com; https://john.turri.org/
laurie.santos@yale.edu; https://caplab.yale.edu/
joshua.knobe@yale.edu; https://campuspress.yale.edu/joshuaknobe/

Target Article

Cite this article: Phillips J, Buckwalter W, Cushman F, Friedman O, Martin A, Turri J, Santos L, Knobe J. (2021) Knowledge before belief. *Behavioral and Brain Sciences* **44**, e140: 1–75. doi:10.1017/S0140525X20000618

Target Article Accepted: 03 September 2020
Target Article Manuscript Online: 08 September 2020
Commentaries Accepted: 30 December 2021

Key words:
belief; factivity; false belief; knowledge; knowledge-first; theory of mind

What is Open Peer Commentary? What follows on these pages is known as a Treatment, in which a significant and controversial Target Article is published along with Commentaries (p. 17) and an Author's Response (p. 67). See bbsonline.org for more information.

Abstract

Research on the capacity to understand others' minds has tended to focus on representations of *beliefs*, which are widely taken to be among the most central and basic theory of mind representations. Representations of *knowledge*, by contrast, have received comparatively little attention and have often been understood as depending on prior representations of belief. After all, how could one represent someone as knowing something if one does not even represent them as believing it? Drawing on a wide range of methods across cognitive science, we ask whether belief or knowledge is the more basic kind of representation. The evidence indicates that nonhuman primates attribute knowledge but not belief, that knowledge representations arise earlier in human development than belief representations, that the capacity to represent knowledge may remain intact in patient populations even when belief representation is disrupted, that knowledge (but not belief) attributions are likely automatic, and that explicit knowledge attributions are made more quickly than equivalent belief attributions. Critically, the theory of mind representations uncovered by these various methods exhibits a set of signature features clearly indicative of knowledge: they are not modality-specific, they are factive, they are not just true belief, and they allow for representations of egocentric ignorance. We argue that these signature features elucidate the primary function of knowledge representation: facilitating learning from others about the external world. This suggests a new way of understanding theory of mind – one that is focused on understanding others' minds in relation to the actual world, rather than independent from it.

1. Introduction

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Mindreading in conversation

Evan Westra [a,*], Jennifer Nagel [b]

[a] *Department of Philosophy, York University, Canada*
[b] *Department of Philosophy, University of Toronto, Canada*

ARTICLE INFO

Keywords:
Mentalizing
Conversation
Factivity

ABSTRACT

How is human social intelligence engaged in the course of ordinary conversation? Standard models of conversation hold that language production and comprehension are guided by constant, rapid inferences about what other agents have in mind. However, the idea that mindreading is a pervasive feature of conversation is challenged by a large body of evidence suggesting that mental state attribution is slow and taxing, at least when it

DOI: 10.1111/phpr.70016

Philosophy and Phenomenological Research

ORIGINAL ARTICLE

Assertoric mindreading

Peter van Elswyk

Department of Philosophy, Northwestern University, Evanston, Illinois, USA

Correspondence
Peter van Elswyk, Department of Philosophy, Northwestern University, Evanston, IL 60208, USA.
Email: peter.vanelswyk@gmail.com

Abstract

This essay offers an explanation of how assertions express that the speaker has a propositional attitude toward what's asserted. The explanation is that this feature of assertion is owed to a hearer's spontaneous mindreading. I call this the ASSERTORIC MINDREADING HYPOTHESIS. Once developed and defended, the hypothesis is used to investigate which attitude is expressed. Since the attitude expressed is the attitude tracked during mindreading, the attitude must have a certain profile. It is argued that only factive attitudes like knowledge have this profile. Non-factive attitudes like belief or acceptance are ineligible.

Knowledge Attribution

Belief Attribution

develops earlier than

Children acquire the ability to implicitly attribute knowledge early

Children acquire the ability to implicitly attribute belief late (and the relevant experiments don't replicate)

Knowledge Attribution

develops earlier than

Belief Attribution

Children acquire the ability to implicitly attribute knowledge early

Children acquire the ability to implicitly attribute belief late (and the relevant experiments don't replicate)

Children acquire the ability to verbally attribute "know" early

Children acquire the ability to verbally attribute "think" and "believe" late

Knowledge Attribution

Belief Attribution

is faster and less effortful than

Knowledge Attribution

is faster and less effortful than

Belief Attribution

Children's performance an executive function tasks is not correlated with their performance on knowledge tasks

Children's performance an executive function tasks is *predicted* by their performance on (false) belief tasks

Knowledge Attribution

is faster and less effortful than

Belief Attribution

Children's performance an executive function tasks is not correlated with their performance on knowledge tasks

Does not effortfully rely on working memory, inhibition control, or cognitive flexibility for task performance

Children's performance an executive function tasks is *predicted* by their performance on (false) belief tasks

Effortfully relies on working memory, inhibition control, or cognitive flexibility for task performance

WHEN IS KNOWLEDGE (AS OPPOSED TO BELIEF) ATTRIBUTED IN CONVERSATION?

WHEN IS KNOWLEDGE (AS OPPOSED TO BELIEF) ATTRIBUTED IN CONVERSATION?

Often! Hard to know exactly, though, as many of the relevant experiments in psycholinguistics aren't set-up to distinguish between factive and non-factive mindreading.

Attitude verbs

Children initially mislearn “think” as a veridical verb like “know” because they are able to attribute knowledge as opposed to belief

Declaratives and interrogatives

In response to the use of a declaratives and interrogatives, hearers attribute knowledge (or ignorance) to the speaker by default.

Declaratives and interrogatives

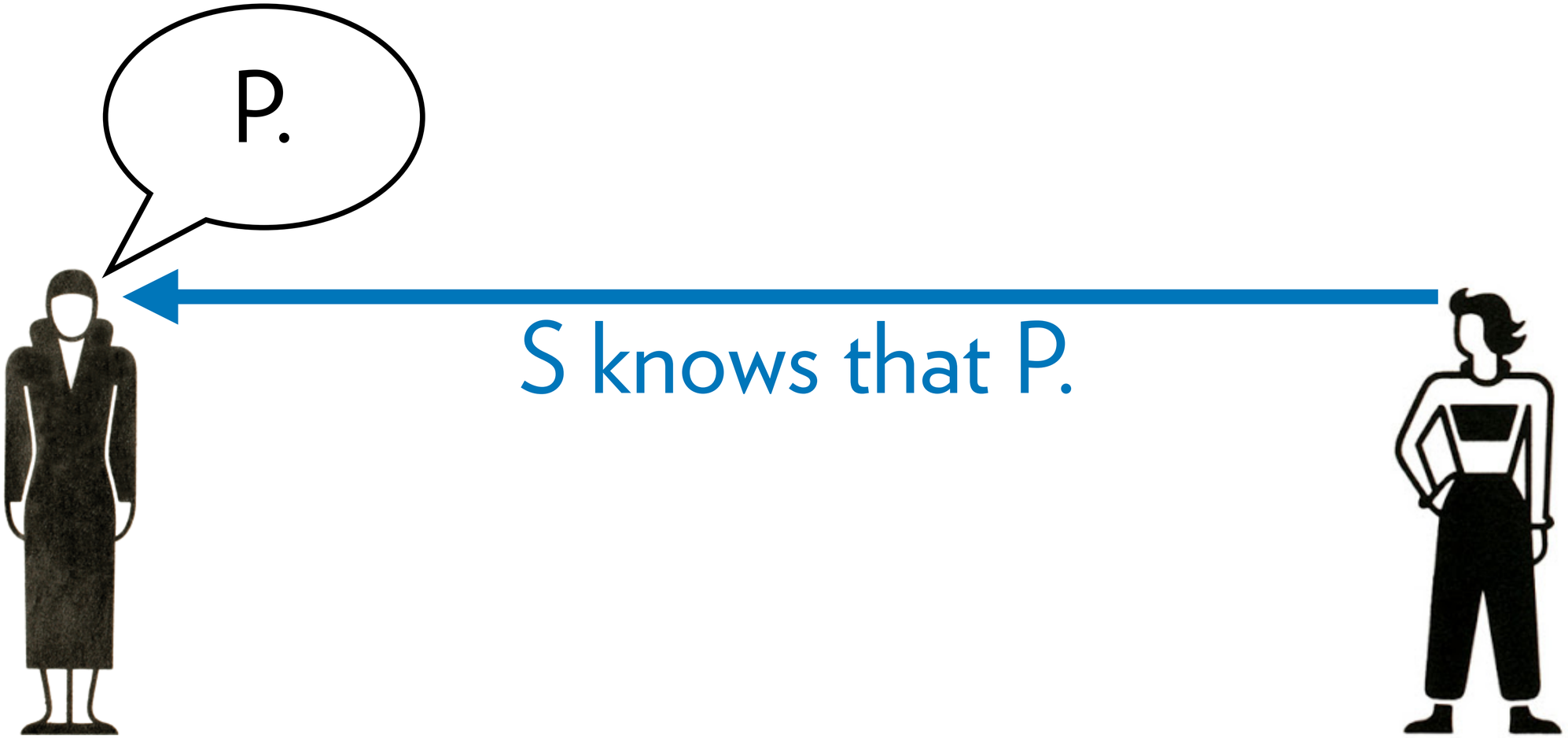
In response to the use of declaratives and interrogatives, hearers attribute knowledge to the speaker by default.

Declaratives and interrogatives

In response to the use of declaratives and interrogatives, hearers attribute knowledge to the speaker by default.

Declaratives and interrogatives

In response to the use of declaratives and interrogatives, hearers attribute knowledge to the speaker by default.

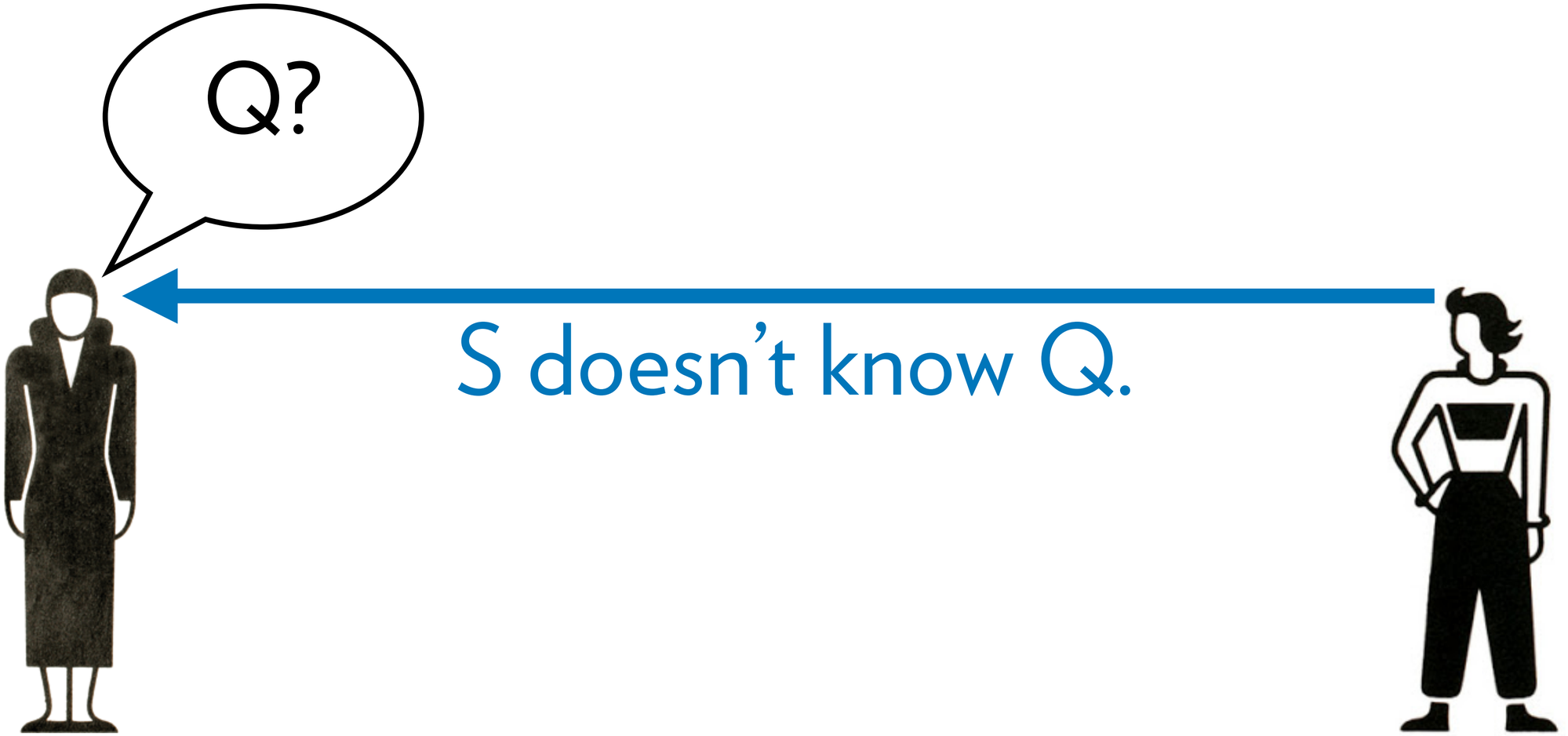


Declaratives and interrogatives

In response to the use of declaratives and interrogatives, hearers attribute knowledge to the speaker by default.

Declaratives and interrogatives

In response to the use of declaratives and interrogatives, hearers attribute knowledge to the speaker by default.



ELSEVIER

Cognition

journal homepage: www.elsevier.com/locate/cognit

COGNITION

Original Articles

How do you know that? Automatic belief inferences in passing conversation

Paula Rubio-Fernández[a,b,*], Francis Mollica[c], Michelle Oraa Ali[a], Edward Gibs

[a] Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, United States
[b] Department of Philosophy, University of Oslo, Norway
[c] Department of Brain and Cognitive Sciences, University of Rochester, United States

ARTICLE INFO

Keywords:
Theory of Mind
False-belief tasks
Automatic and controlled processes
Pragmatics
Belief inferences

ABSTRACT

There is an ongoing debate, both in philosophy and psychology
infer what others may know, or whether they can only derive be
Evidence from laboratory tasks, often involving false beliefs c
belief inferences are cognitively costly, controlled processes. H
belief reasoning is pervasive and therefore potentially automa
conducted two pre-registered self-paced reading experiments (
ments showed that participants slowed down when a stranger co
relative to conditions where a stranger commented on their own
which violated participants' common-ground expectations. We
account for belief reasoning in conversation as it is at the cent

ELSEVIER

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

COGNITIO

How do we interpret questions? Simplified representations of knowledge guide humans' interpretation of information requests

Marie Aguirre[a,*], Mélanie Brun[a], Anne Reboul[b], Olivier Mascaro[a]

[a] Université de Paris, INCC UMR 8002, CNRS, F-75006 Paris, France
[b] Laboratory of Cognitive Psychology, UMR 7290, CNRS and Aix-Marseille University, Marseille, France

ARTICLE INFO

Keywords:
Theory of mind
Questions
Knowledge
Pragmatics
Cognitive development
Naïve epistemology

ABSTRACT

This paper investigates the cognitive mechanisms supporting humans' interpretation of requests for informatio
Learners can only search for a piece of information if they know that they are ignorant about it. Thus,
principle, the interpretation of requests for information could be guided by representations of Socratic ignoran
(tracking what people know that they do not know). Alternatively, the interpretation of requests for informati
could be simplified by relying primarily on simple knowledge tracking (i.e., merely tracking what people know
We judged these hypotheses by testing two-and-a-half-year-old toddlers ($N = 18$), five- to seven-year-old childr
($N = 72$), and adults ($N = 384$). In our experiments, a speaker asked a question that could be disambiguated
tracking her state of knowledge. We manipulated the speakers' visuals to modulate the complexity of t
ignorance representation required to disambiguate their questions. Toddlers showed no tendency to appeal
representations of Socratic ignorance when disambiguating questions (Pilot S1). Five- to seven-year-ol
exhibited a similar pattern of results, and they performed better when information requests could be disa
biguated using simple knowledge tracking (Studies 1a-1b). Adults used representations of Socratic ignorance
interpret questions, but were more confident when simple knowledge tracking was sufficient to disambigua
information requests (Studies 2–3). Moreover, adults disambiguated questions as if speakers could request i
formation about things that they were ignorant of, even when speakers had no reason to know about the

A cognitive default

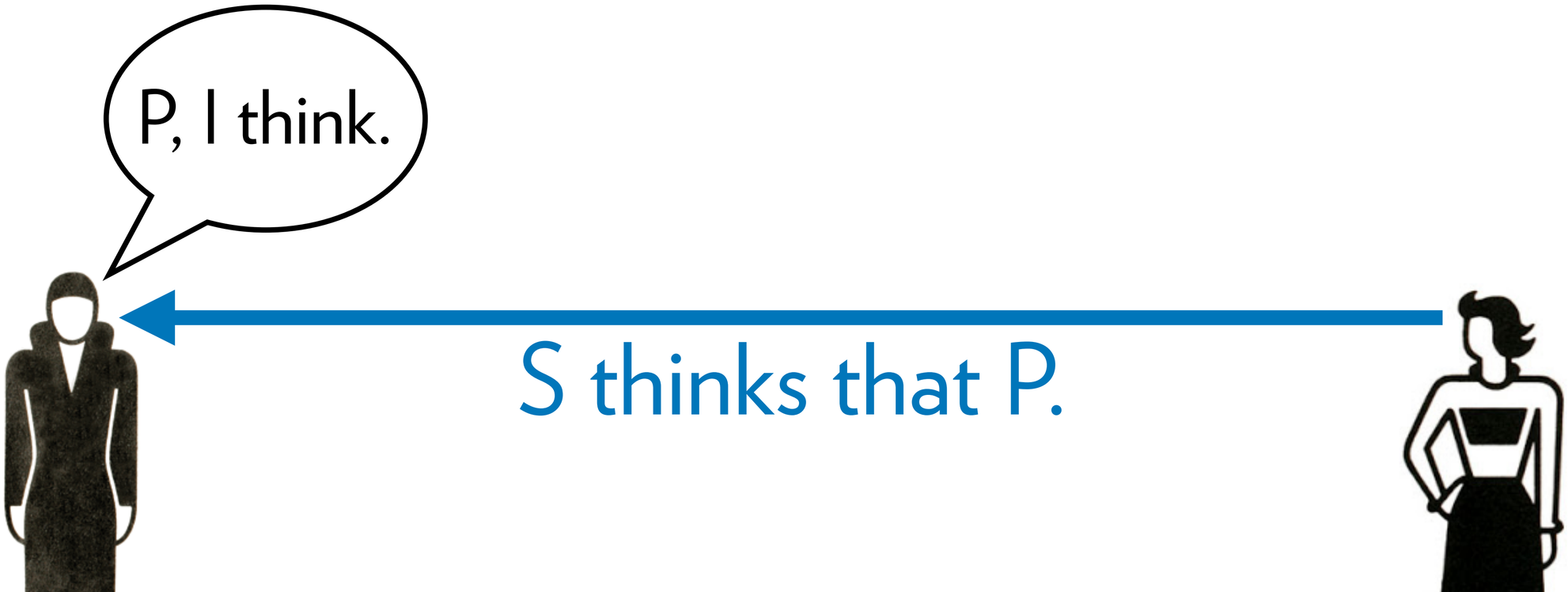
The attribution of knowledge (or ignorance) is a default cued by the use of an unmarked sentence. Marked interrogatives or declaratives cue different forms of mindreading.

A cognitive default

The attribution of knowledge (or ignorance) is a default cued by the use of an unmarked sentence. Marked interrogatives or declaratives cue different forms of mindreading.

A cognitive default

The attribution of knowledge (or ignorance) is a default cued by the use of an unmarked sentence. Marked interrogatives or declaratives cue different forms of mindreading.



What this explains

- Why Moorean conjunctions like “P, I but I don’t know P” are infelicitous with bare declaratives but felicitous with hedged declaratives
- Why bare declaratives can be challenged with “How do you know that?” but hedged declaratives cannot be
- Why bare declaratives are prompted by bare interrogatives like “Q?” but hedged declaratives are prompted by marked interrogatives like “Q, do you think?”

WHEN IS KNOWLEDGE (AS OPPOSED TO BELIEF) ATTRIBUTED IN CONVERSATION?

Often! Hard to know exactly, though, as many of the relevant experiments in psycholinguistics aren't set-up to distinguish between factive and non-factive mindreading.

Often! Hard to know exactly, though, as many of the relevant experiments in psycholinguistics aren't set-up to distinguish between factive and non-factive mindreading.

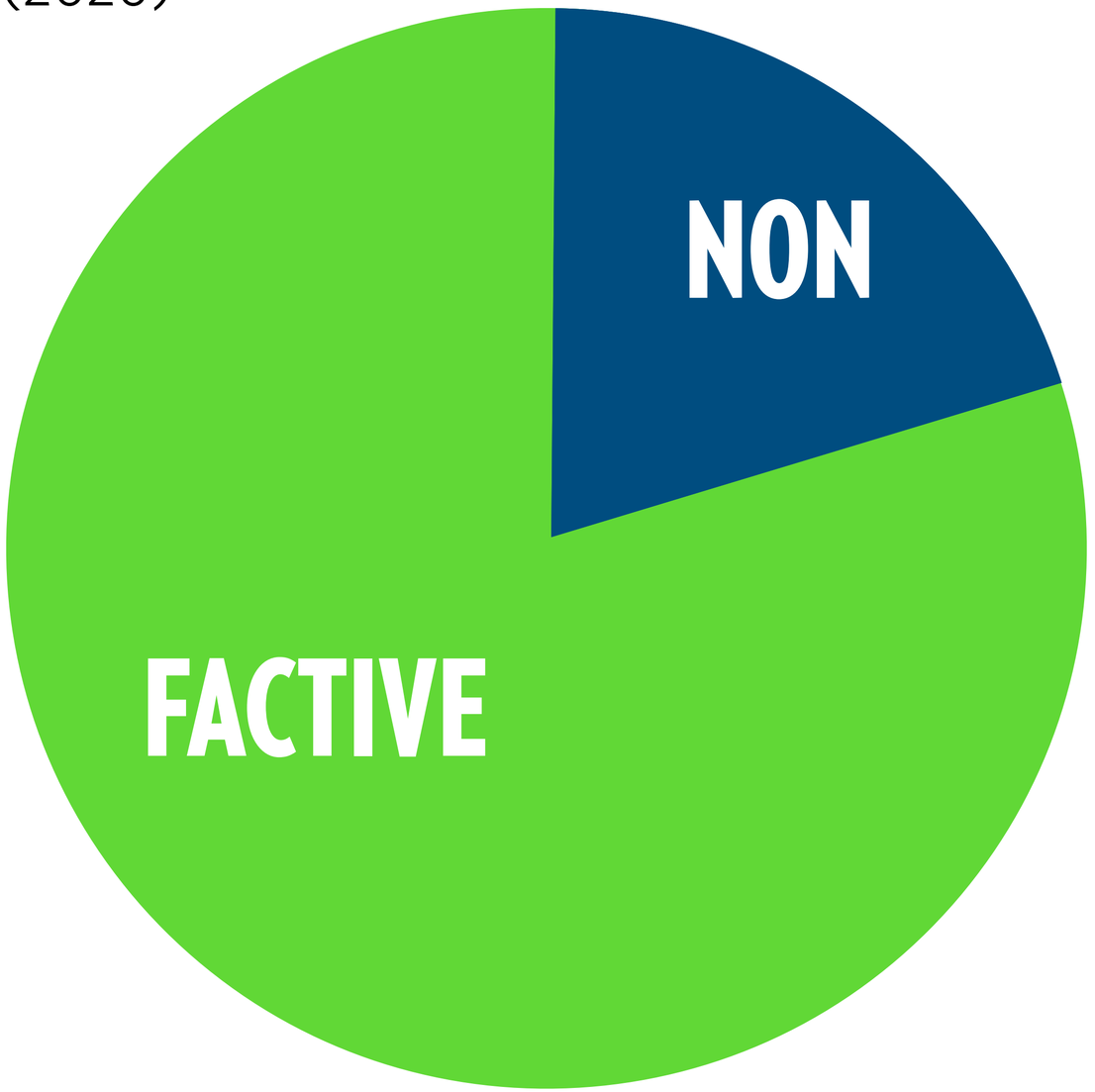
Often! Hard to know exactly, though, as many of the relevant experiments in psycholinguistics aren't set-up to distinguish between factive and non-factive mindreading.

- Maybe reference tasks are explained by tracking knowledge and ignorance?
- Maybe attentional tasks are explained by tracking what a person knows about through sight or what they are positioned to know?
- Maybe...

When do we need non-factive mindreading?

Westra and Nagel (2020)

- Deliberate deception
- Epistemic vigilance
- Metalinguistic repair
- Argumentation



DISCOURSE PROCESSES
2020, VOL. 57, NOS. 5–6, 441–459
https://doi.org/10.1080/0163853X.2020.1719795

Conversational Repair and the Acquisition of Language

Eve V. Clark

Department of Linguistics, Stanford University

ABSTRACT

In this article, I examine how repairs in adult-child conversations guide children's acquisition of language. Children make unprompted self-repairs to their utterances. They also respond to prompts for repair, whether open (*Hm?, What?*) or restricted (*You hid what?*), and to restricted offers (Child: *I falled*, Adult: *You fell?*). Children respond to clarification requests with self-repairs in the next turn, and make use of the feedback offered. The contrast between their utterance and the adult utterance identifies the locus of the error (negative feedback), while the adult's offer presents a conventional version of the child's utterance (positive feedback). I describe the use of restricted offers in conversations with children acquiring English and French, then present two case studies of how these inform children about homophonous French verb forms and early opaque Hebrew verb uses. These findings demonstrate the fundamental role of repair in the acquisition of a first language.

Introduction

Children learn language from the speakers around them, in the to-and-fro of conversational exchanges (Clark, 2018a), but they start out with minimal knowledge about how to do things with words, and so may display difficulty in producing the terms needed to convey their intentions in a comprehensible form to their addressees. One result is that they frequently need to repair their own speech. For example, Golinkoff (1986) found that about 50% of infant-initiated interactions (aged 1;0–1;5) involved other-initiated repairs.

They are also responsive to requests for repair from others. Both self-initiated and other-initiated repairs, I argue, play a basic role in the process of acquisition. In essence, they offer a way of giving children feedback – in particular, negative feedback when they have made an error, often combined with positive feedback when adults provide a conventional way to express a particular meaning. Repairs offer one resource for adults offering children feedback during acquisition, and adults do offer feedback, despite past claims to the contrary.[1]

The Social Brain Hypothesis and Human Evolution

Robin I. M. Dunbar

https://doi.org/10.1093/acrefore/9780190236557.013.44
Published online: 03 March 2016

Summary

Primate societies are unusually complex compared to those of other animals, and the need to manage such complexity is the main explanation for the fact that primates have unusually large brains. Primate sociality is based on bonded

CellPress

Trends in Cognitive Sciences

Review

Interactive repair and the foundations of language

Mark Dingemanse [1,*] and N.J. Enfield[2]

The robustness and flexibility of human language is underpinned by a machinery of interactive repair. Repair is deeply intertwined with two core properties of human language: reflexivity (it can communicate about itself) and accountability (it is used to publicly enforce social norms). We review empirical and theoretical advances from across the cognitive sciences that mark interactive repair as a domain of pragmatic universals, a key place to study metacognition in interaction, and a system that enables collective computation. This provides novel insights into the role of repair in comparative cognition, language development, and human–computer interaction. As an always-available fallback option and an infra-

Highlights

A frequently used system of interactive repair scaffolds the complexity of human language.

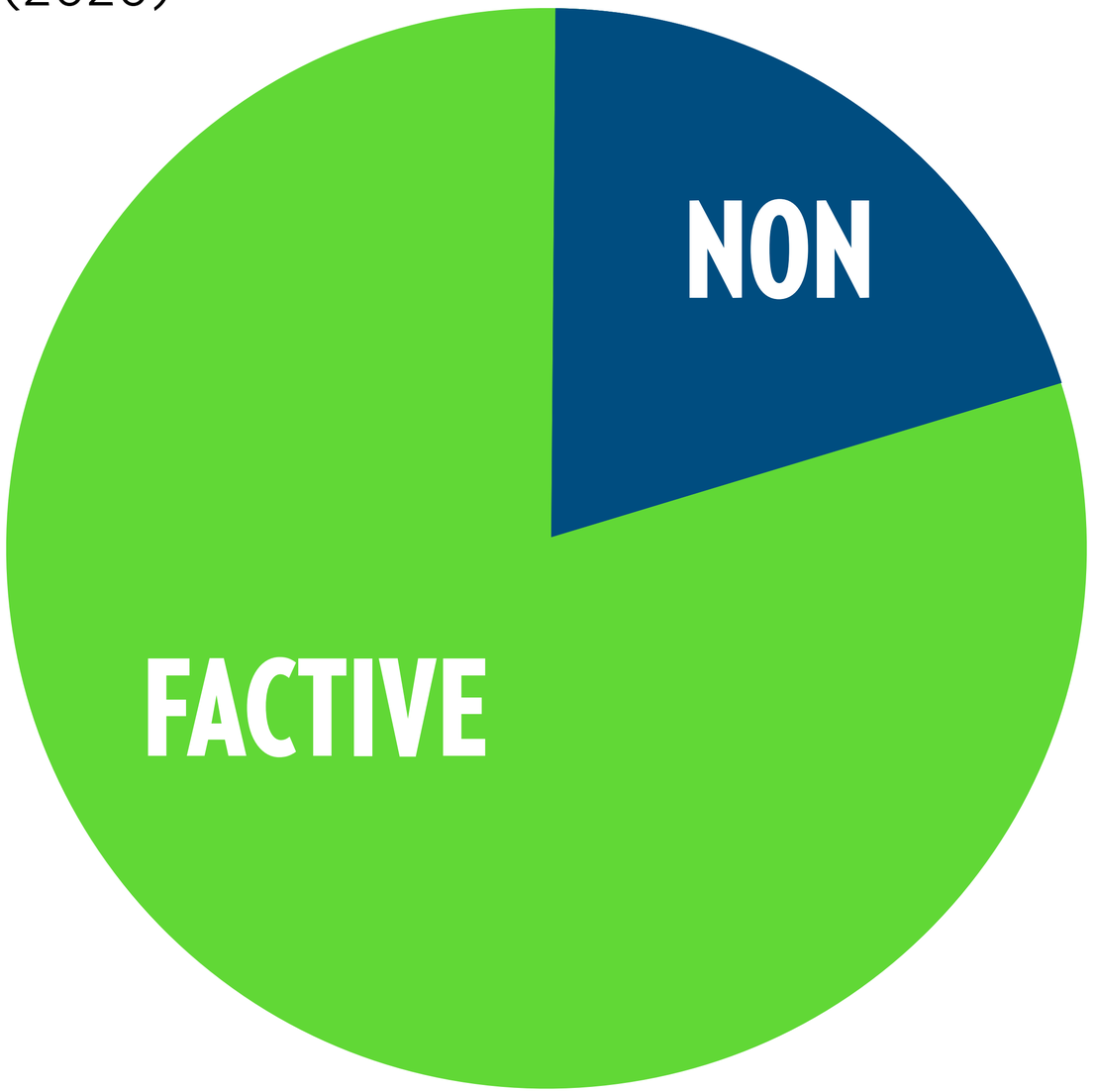
Interactive repair is both a mechanism for ensuring informational robustness and an organisation for social accountability.

Repair embodies and exploits the unique reflexivity of human language.

When do we need non-factive mindreading?

Westra and Nagel (2020)

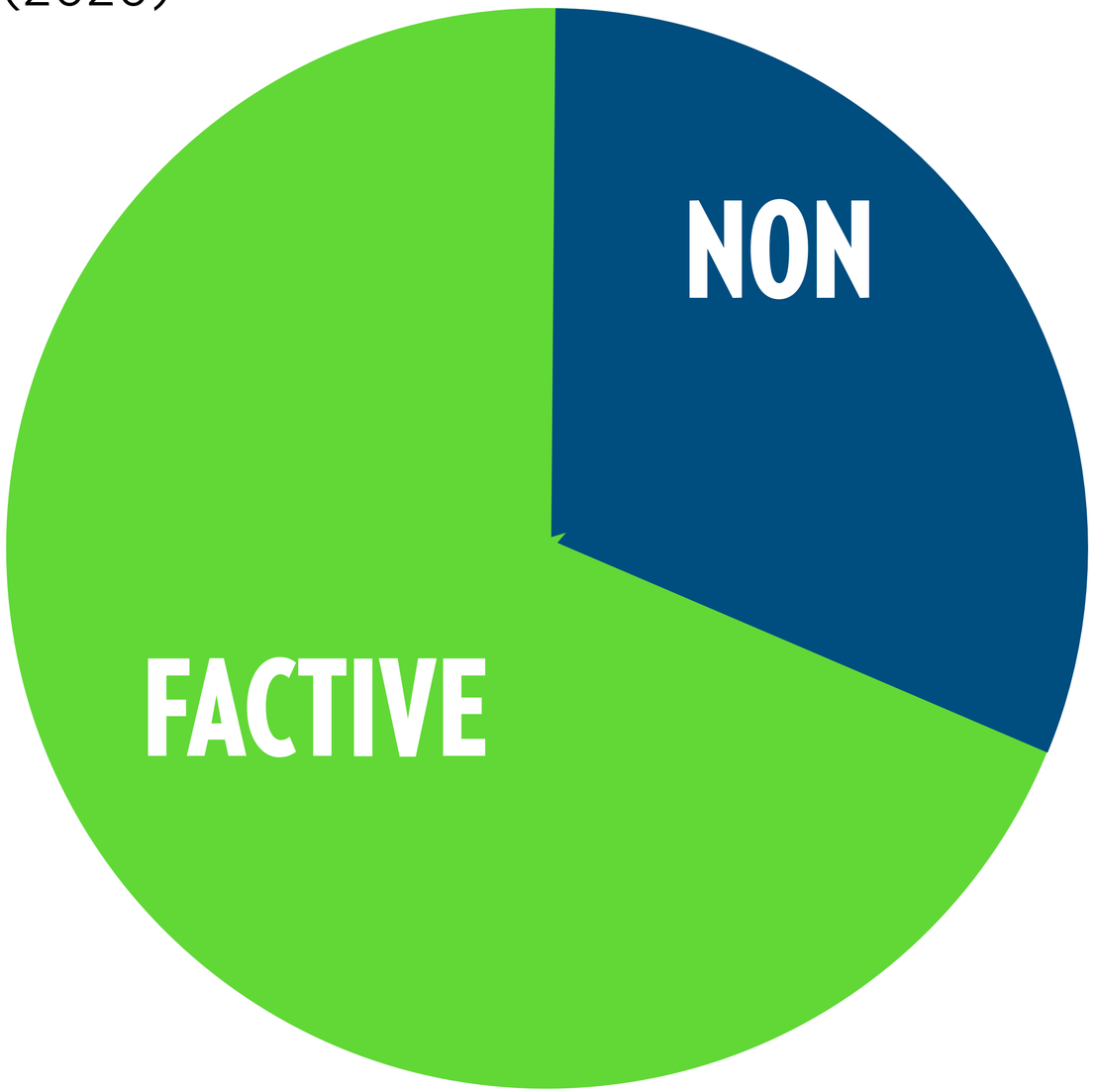
- Deliberate deception
- Epistemic vigilance
- Metalinguistic repair
- Argumentation



When do we need non-factive mindreading?

Westra and Nagel (2020)

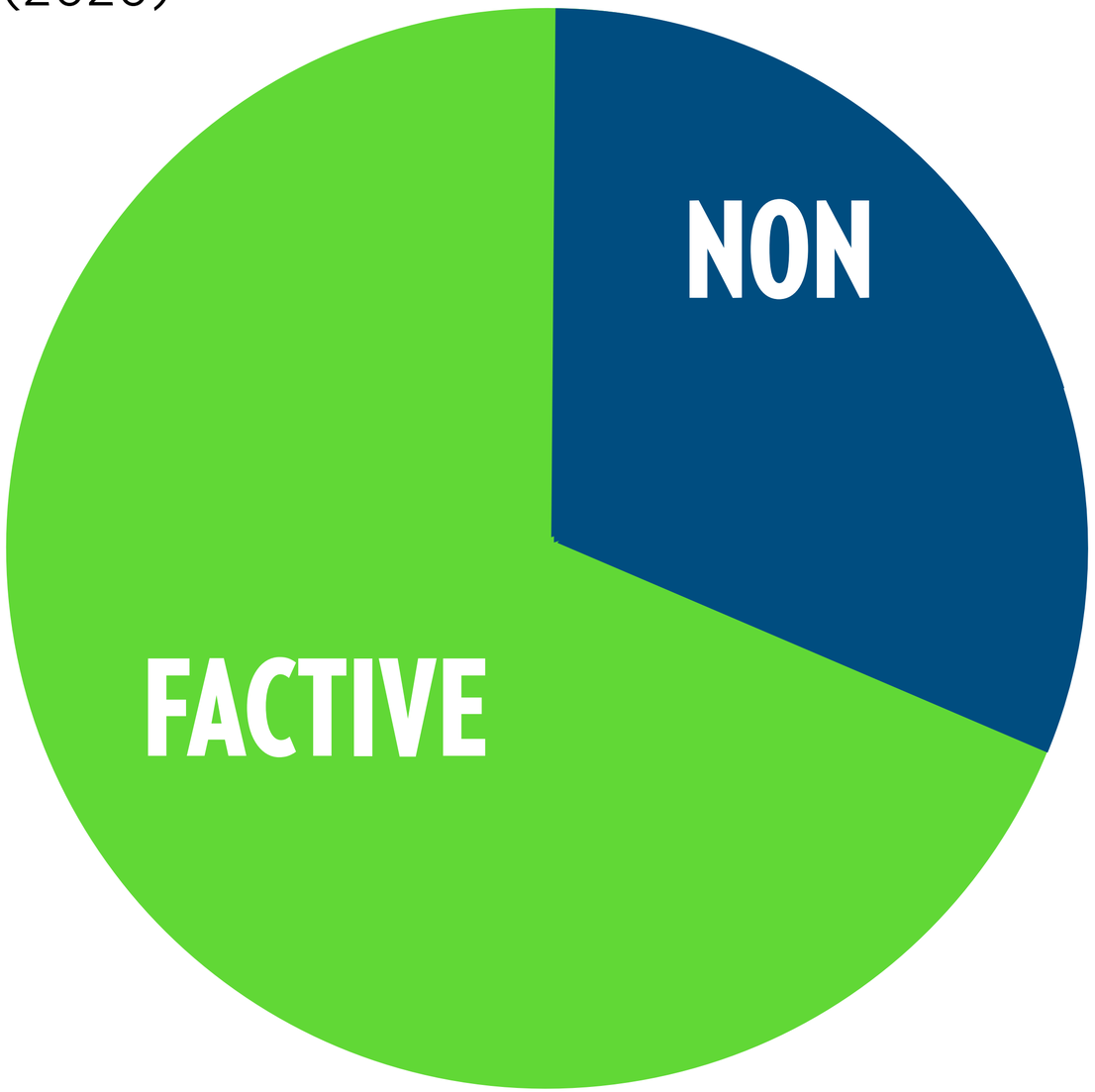
- Deliberate deception
- Epistemic vigilance
- Metalinguistic repair
- Argumentation



When do we need non-factive mindreading?

Westra and Nagel (2020)

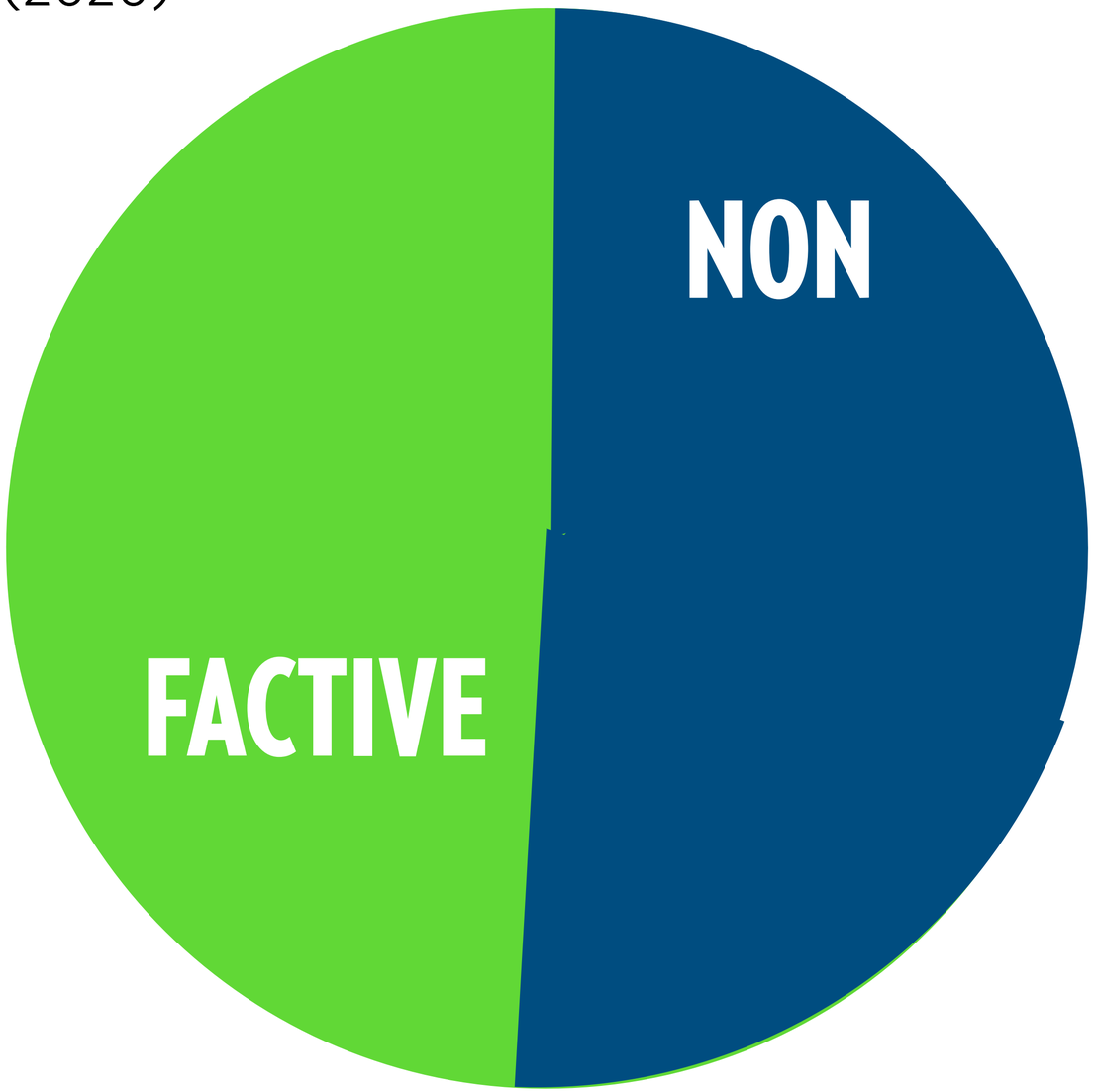
- Deliberate deception
- Epistemic vigilance
- Metalinguistic repair
- Argumentation
- Avoiding disagreements
- Pretense
- Predictions



When do we need non-factive mindreading?

Westra and Nagel (2020)

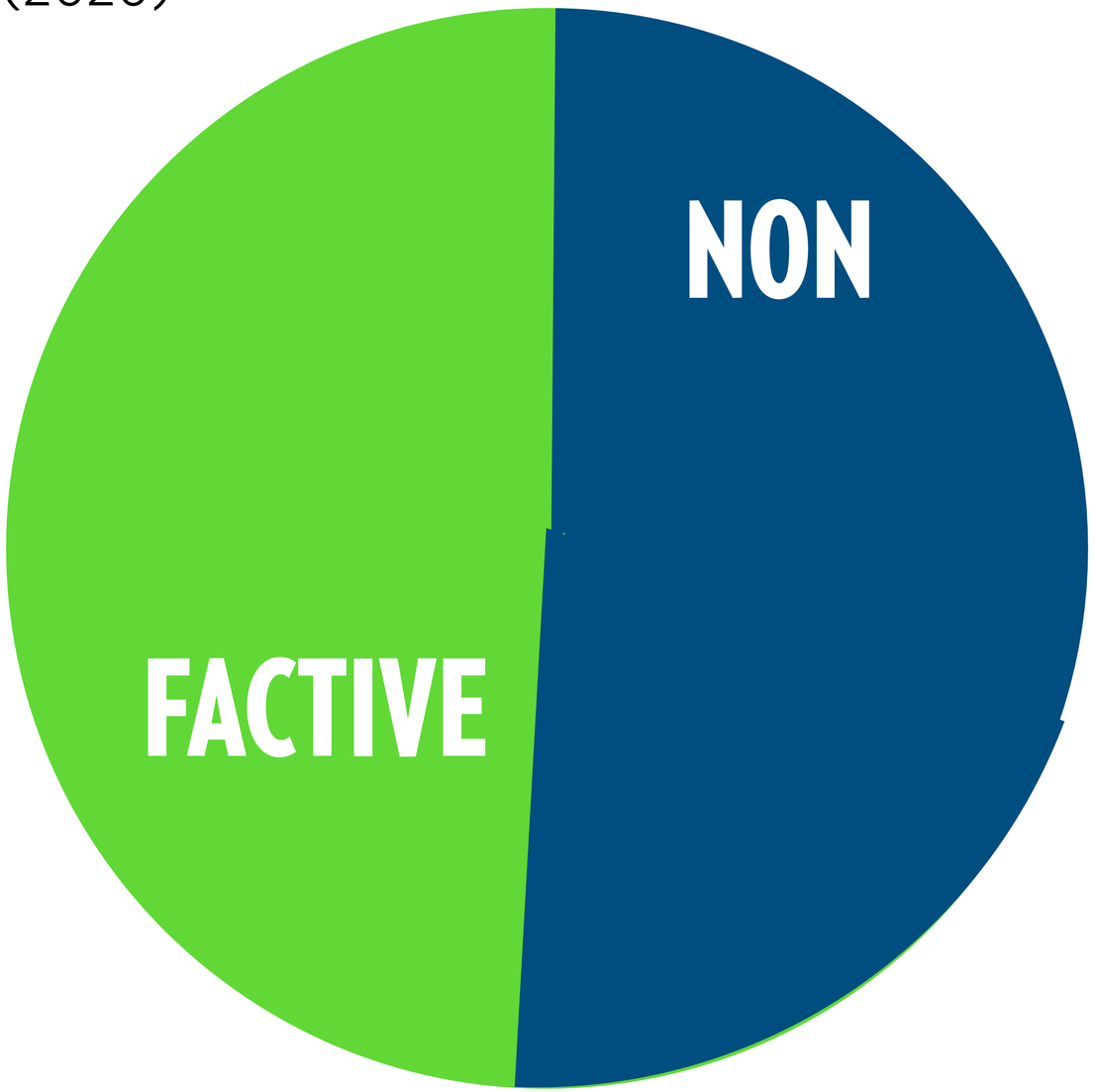
- Deliberate deception
- Epistemic vigilance
- Metalinguistic repair
- Argumentation
- Avoiding disagreements
- Pretense
- Predictions



When do we need non-factive mindreading?

Westra and Nagel (2020)

- Intention/Plan Attribution
- Desire Attribution
- Object Representation Attribution
- Partial Belief/credence attribution
- ⋮



When do we need non-factive mindreading?

Westra and Nagel (2020)

- Intention/Plan Attribution
- Desire Attribution
- Object Representation Attribution
- Partial Belief/credence attribution
- ⋮