



Cognitive Science 47 (2023) e13376

© 2023 The Authors. *Cognitive Science* published by Wiley Periodicals LLC on behalf of *Cognitive Science Society (CSS)*.

ISSN: 1551-6709 online

DOI: 10.1111/cogs.13376

Against Arguments From Diagnostic Reasoning

Jeske Toorman

Department of Philosophy and Religious Studies, Norwegian University of Science and Technology

Received 1 May 2023; received in revised form 30 August 2023; accepted 27 October 2023

Abstract

Recent work in cognitive psychology and experimental semantics indicates that people do not categorize natural kinds solely by virtue of their purported scientific essence. Two attempts have been made to explain away the data by appealing to the idea that participants in these studies are reasoning diagnostically. I will argue that an appeal to diagnostic reasoning will likely not help to explain away the data.

Keywords: Experimental semantics; Psychological essentialism; Categorization; Concepts; Teleological essentialism; Diagnostic reasoning

1. Introduction

According to a well-established view in both cognitive psychology and the philosophy of language, a substance is to be categorized as an instance of a natural kind if and only if it shares the presumed scientific essence responsible for the superficial features of that natural kind (e.g., Gelman, 2003; Kripke, 1980; Putnam, 1975). Recently, this “scientific essence view” has been challenged by data indicating that either superficial properties generally or an entity’s Aristotelean telos specifically guide categorization behavior (Haukioja, Nyquist, & Jylkkä, 2021; Rose & Nichols, 2019). Proponents of the scientific essence view have responded with what I will call “arguments from diagnostic reasoning” (Devitt & Porter, 2021; Neufeld, 2021). According to arguments from diagnostic reasoning, participants in

Correspondence should be sent to Jeske Toorman, Department of Philosophy and Religious Studies, Norwegian University of Science and Technology, 7491 Trondheim, Norway. E-mail: Jeske.toorman@ntnu.no

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

these studies base their answer on an inference from external features to scientific essence, in such a way that their response is compatible with the scientific essence view.

In this paper, I will argue that an appeal to diagnostic reasoning is unlikely to help the scientific essence view. Section 1 will briefly elaborate on the scientific essence view, its recent challenges, and the responses that make an appeal to arguments from diagnostic reasoning. Section 2 will argue against appealing to diagnostic reasoning to defend the scientific essence view. Section 3 will address a recent study by Joo and Yousif (2022), purporting to provide empirical evidence for the claim that people reason diagnostically in a way that supports the scientific essence view.

2. The scientific essence view and arguments from diagnostic reasoning

The scientific essence view, as I will understand it here, is any view according to which the only relevant criterion for the categorization of a natural kind is the deep property responsible for the superficial features of the members of that kind. In cognitive psychology, this view goes by the name *psychological essentialism*, and it is a view according to which people believe that natural kinds have “an underlying reality or true nature, shared by members of a category, that one cannot observe directly but that gives an object its identity and is responsible for the other similarities that category members share” (Gelman, 2003, p. 7–8). In its placeholder variety, these essence beliefs are general beliefs. That is, the relevant essence belief is not, for example, that the essence of water is being composed of H₂O, or that the essence of horses is having a certain kind of DNA. Rather, they are beliefs to the effect that natural kinds have an essence, and that essence just is whatever gives rise to the observable similarities.

In the philosophy of language, the most well-known instance of the view goes by the name *physical externalism*, and it is a view according to which the reference of natural kind terms supervenes on the deep properties, whatever they might be, that are responsible for the observable similarities of most of the substances to which the relevant term has been applied (e.g., Putnam, 1975).

In both cases, what is responsible for the observable similarities among members of a kind is typically assumed to be a property or set of properties discoverable by the natural sciences, as opposed to the kind of properties easily perceivable by the layman. Further, although the former view is typically taken to be a view about people’s beliefs or narrowly construed concepts, whereas the latter view is a view about language, both views are taken to be supported or undermined simultaneously by means of the same categorization behavior. As the concern of this paper is with the extent to which categorization behavior supports both of these accounts, I will treat them here as a single view.

In the philosophy of language, the scientific essence view has recently been challenged by Haukioja et al. (2021). Haukioja et al. presented participants with two types of scenarios. In the first type, the Twin Earth cases, space travelers find a substance on a planet in a different galaxy that looks like an instance of a familiar kind but has a chemical composition different from that kind. In the second type, the reverse Twin Earth cases, space travelers find

a substance that has a chemical composition identical to that of a familiar kind, but which nevertheless looks and acts radically different than instances of the familiar kind. When participants had to answer a forced-choice question, most of them chose an answer indicating that they took the substance in neither scenario to be an instance of the familiar kind. However, assuming participants take chemical composition to be responsible for external features, they should, according to the scientific essence view, categorize the substance in the latter scenario to be an instance of the familiar kind.¹

In cognitive psychology, the scientific essence view has recently been challenged by Rose and Nichols (2019). Rose and Nichols (henceforth: R&N) presented subjects with two experiments. In their first study, R&N presented subjects with a case in which a bee is made to look like a spider. Subsequently, in one condition, the animal in question retained the purported telos of a bee (viz., pollinating flowers), whereas in another condition, the animal acquires the purported telos of a spider (viz., spinning webs). In a second study, R&N presented subjects with a case in which the insides of a bee are removed and replaced by the insides of a spider. Again, in one condition, the animal in question retains the telos of a bee, whereas in another condition, it acquires the telos of a spider. In both studies, subjects judged the animal to be more like a bee than a spider when it retained the telos of a bee, and more like a spider when it acquired the telos of a spider. Both studies suggest that neither scientific essence nor superficial properties are necessary for being categorized as an instance of a natural kind. The telos, on the other hand, does seem to be.² Hence, R&N dub this view “teleological essentialism.”

Both Haukioja et al. and R&N have been criticized for not taking into account that participants might reason diagnostically in their experiments. People reason diagnostically when they make inferences concerning the causes of effects on the basis of knowledge of these effects. In Haukioja et al.’s study, subjects fail to categorize as “water” a substance that is composed of H₂O, but which does not look like the stuff commonly known as “water.” Now, as Devitt and Porter (2021) point out, according to the scientific essence view, the scientific essence, whatever it may be, is that which causes the observable features of a kind. Hence, if there are two substances looking and behaving unlike one another but which share being composed of H₂O, it follows that being composed of H₂O is unlikely to be the scientific essence of at least one of these substances. But not categorizing a substance as “water” because it is assumed not to have the scientific essence of the stuff commonly known as “water” is of course exactly what the scientific essence view predicts. Devitt and Porter conclude that “Kripke-Putnam [i.e., the scientific essence view] can explain the reverse TE [i.e., Twin Earth] results” (p. 29).

In a similar vein, Neufeld (2021) points out that R&N’s study involves the manipulation of those features that are supposed to be caused by the scientific essence. These features, however, allow participants to infer the scientific essence, and from the scientific essence, category membership can be inferred. So, according to Neufeld, a bee that is made to look like a spider and has acquired the telos of a spider is judged to be a spider not because it has the telos of a spider but rather because having the telos of a spider is taken to be diagnostic evidence of having the scientific essence of a spider. Similarly, a bee that has been given the insides of a spider but still has the telos of a bee is judged to be a bee not because it has the

telos of a bee but because it is inferred to have the scientific essence of a bee. This, however, is just what the scientific essence view predicts. Hence, according to Neufeld, the data presented by R&N do not support teleological essentialism over the scientific essence view. “To the contrary, their findings are straightforwardly predicted by SEH [i.e., the scientific essence view]” (p. 2).

It should be noted that, in addition to Haukioja et al.’s and R&N’s study, there are more studies indicating that people do not, or not solely, categorize natural kinds by virtue of scientific essence (see, e.g., Braisby, Franks, & Hampton, 1996; Hampton, Estes, & Simmons, 2007; Machery et al., 2023; Malt, 1994; Tobia, Newman, & Knobe, 2020). For at least some of these studies, one might very well construct an argument from diagnostic reasoning to attempt to explain the results as being in accordance with the scientific essence view. However, the focus of this paper will be on the study by Haukioja et al. and R&N because only they have been actual targets of arguments from diagnostic reasoning.

3. Against arguments from diagnostic reasoning

This section will provide three arguments to the effect that an appeal to diagnostic reasoning is unlikely to help the scientific essence view. The first argument is this. R&N’s first experiment is inspired by a study by Keil (1989), in which an animal is made to look and act like another animal. In line with psychological essentialism, many subjects in Keil’s study answer that the animal in question still belonged to its original category. Assuming that the scientific essence view is true, it follows that participants in Keil’s study do not reason diagnostically from a change in appearance and behavior to a change in scientific essence. For they answer that the animal is of the old category, whereas it has the outside appearance and behavior of the new category. Consequently, for the scientific essence view to be able to explain away R&N’s findings, there must be a difference between Keil’s study on the hand and that of R&N on the other, which is such that participants reason diagnostically in the latter study, but not in the former. In the absence of such a difference, one might conclude from the fact that no inference is being made in the former case that no inference is being made in the latter case either.

According to Neufeld, there is indeed a relevant difference. In Keil’s study, participants are told that the outward appearance and behavior are caused by an external intervention, whereas in R&N’s study, they are not. This is crucial, according to Neufeld, because:

[W]hen we know that the effect features have been generated by external background causes and not the causal essence, the inference from effect features to underlying essence is defeated. Hence, effects that normally provide evidence for underlying features can be explained away. (p. 4)

However, the actual differences between the two cases are unlikely to make for a difference in diagnostic reasoning. In both R&N’s first experiment and the study of Keil, the change in appearance is explicitly said to be due to a “special operation,” and in both experiments,

the change in behavior is explicitly said to have come about due to training, which is exactly another external intervention. The only thing that is not explicitly said to have come about due to training in R&N's case is the fact that the animal in question spins webs to catch insects. At the same time, it is not unreasonable to suppose that participants take this to be implied by the claim in R&N's vignette stating that the scientists "inserted into the back of it something for making webs" and that they "trained the animal so that it would eat insects." Additionally, after receiving training to eat insects and after having something inserted on its back for making webs, it is not unreasonable to suppose that an animal starts to spin webs to catch insects on its own, without thereby also having a different scientific essence.

Now, a proponent of the scientific essence view might try to appeal to the idea that adults are more likely to reason diagnostically, since Keil's study involved children, whereas R&N's did not. However, there is some evidence that by age 4, almost 6 years younger than the participants in Keil's study, children may already be capable of making diagnostic inferences (Sobel, Yoachim, Gopnik, Meltzoff, & Blumenthal, 2007; see also Fernbach, Macris, & Sobel, 2012). Additionally, using a similar task as used in R&N's first experiment, Rose, Jaramillo, Nichols, and Horne (2022) showed that children younger than the participants in Keil's study already categorize natural kinds by virtue of purported telos. If participants cannot be said to reason diagnostically in Keil's study because of their age, then neither can they be said to reason diagnostically in Rose et al.'s, and hence, these results cannot be explained away.³

For the second argument, assume that participants take what is caused by the scientific essence to be diagnostic evidence of the scientific essence and base their categorization on inferred scientific essence, as arguments from diagnostic reasoning would have it. In that case, it is to be expected that participants would be less likely to categorize an animal as belonging to the category of which it has the telos when telos conflicts with appearance than when telos and appearance match. In the former case, the diagnostic evidence does, after all, conflict.⁴ But participants do not seem to do this. Both R&N's first and second experiments contain a condition in which either the appearance of the animal or the telos of the animal is different from that of the original category, but not both. In their second experiment, participants are indeed slightly less likely to categorize an animal as a spider when it has the appearance of a bee and the purported telos of a spider than they are to categorize it as a bee when it has the appearance and purported telos of a bee. In their first experiment, however, participants are on average slightly *more* likely to categorize the animal as a bee when it looks like a spider and has the telos of a bee than they are to categorize it as a spider when it looks like a spider and has the purported telos of a spider.⁵ It thus seems not to be the case that participants are less likely to infer insides from telos when telos and appearance conflict as compared to when they match.

The third argument I want to make is that an appeal to diagnostic reasoning cannot explain away the results of Haukioja et al. without undermining the appeal to diagnostic reasoning in R&N's experiments.

Haukioja et al. presented participants with two types of vignettes, the Twin Earth cases and the reverse Twin Earth cases. Both types are identical in that participants are told about a newly found substance. In the Twin Earth case specific to "water," this substance has a chemical composition abbreviated as XYZ, but looks and behaves exactly like water as we

are familiar with it. In the reverse Twin Earth cases, by contrast, this substance is composed of H_2O , but looks and behaves very much unlike any kind we are familiar with. Diagnostic reasoning in the way assumed by Devitt and Porter can explain the fact that the found substance in the reverse Twin Earth cases is not categorized as “water” by assuming that participants infer that the found substance does not have the scientific essence of water because it does not look and behave like water. But if appearance and behavior are diagnostic evidence of scientific essence, then so should participants in the standard Twin Earth cases conclude from the fact that the found substance looks and behaves like water that it has the scientific essence of water. In the standard Twin Earth cases, however, participants do not do this, as the found substance is not categorized as water despite looking and behaving like water. But if participants do not reason diagnostically in the Twin Earth cases, why should they be assumed to do so in the reverse Twin Earth cases? After all, the two scenarios differ only with regard to the appearance, behavior, and chemical structure of the found substance.

Defenders of the scientific essence view might want to respond by arguing that diagnostic reasoning is only from *not* looking and behaving like a familiar kind (as the substance in the reverse Twin Earth cases) to *not* having the scientific essence of a familiar kind, but not from looking and behaving like a familiar kind (as the substance in the standard Twin Earth cases) to having the scientific essence of that familiar kind. After all, leaving out the contribution of the external environment, the scientific essence is supposed to be the cause of the superficial features. But whereas causes determine effects, the same effect is compatible with multiple causes.

However, this reply will not help the scientific essence view because the appeal to diagnostic reasoning in R&N’s first and second experiments relies on the very same inference from looking and behaving like a familiar kind to having the scientific essence of that familiar kind. Here too, different scientific essences are compatible with the same external features. Consequently, if that were to be a reason for not taking subjects to reason diagnostically in the Twin Earth cases, then so should it be a reason for not taking them to reason diagnostically in R&N’s first and second experiments.

Notice lastly that in both the standard Twin Earth cases and the reverse Twin Earth cases, it is explicitly stated what the presumed scientific essence (i.e., XYZ or H_2O) of the found substance in question is. As such, one might try to argue that diagnostic reasoning from appearing and behaving like a familiar kind, as the substance in the Twin Earth cases, to having the scientific essence of that familiar kind only goes through when no explicit statement concerning the presumed scientific essence is made. For otherwise, participants would additionally have to take the conclusion of their inference (that the substance is composed of H_2O) to override what they are being told in the vignette (that the substance is composed of XYZ).

Although this might work to save the appeal to diagnostic reasoning to explain the results of R&N’s first experiment, it cannot save the appeal in their second experiment. In R&N’s second experiment, participants are explicitly told that the insides of a bee, and thereby all likely candidates for scientific essence, have been replaced by that of a spider. Nevertheless, the animal is judged to be a bee because, as arguments from diagnostic reasoning would have it, it has the telos of a bee and, therefore, the scientific essence of a bee. This too requires that participants take the conclusion of their inference (that the animal has the scientific essence of a bee) to override what they are being told in the vignette (that the animal has the insides

and thus the scientific essence of a spider).⁶ Hence, if that were to be a reason for not taking subjects to reason diagnostically in Haukioja et al.'s standard Twin Earth cases, so should it be a reason for not taking them to reason diagnostically in R&N's second experiment.

Perhaps there is some other difference that makes it so that people do not reason diagnostically in the Twin Earth cases, but do so in the reverse Twin Earth cases and R&N's experiments. But no other potentially relevant difference is obvious. In the absence of such a difference, it seems that a consistent application of the principles appealed to in order to make arguments from diagnostic reasoning work in Haukioja et al.'s and R&N's experiments shows that participants do not in both cases reason diagnostically in a way that supports the scientific essence view. And of course, one cannot just appeal to diagnostic reasoning when the results do not fit one's preferred theory.

4. Joo and Yousif: Testing diagnostic reasoning in R&N's experiments

Inspired by Neufeld, Joo and Yousif (2022; henceforth: J&Y) aim to provide what they call "stronger test cases of teleological essentialism" (p. 4). Based on their results, they conclude that "the studies reported here results offer [*sic*] empirical support for the argument made by Neufeld" (p. 14). This conclusion is, however, too quick.

In J&Y's first critical experiment, participants are told about two skilled scientists performing a "special experiment" on an animal. Subsequently, they are told explicitly either that the presumed telos of the animal has stayed the same, or that it changed to that of a new category. Additionally, they are explicitly told either that the insides have remained that of the original category, or that they changed to that of an animal of the new category. What J&Y find is that when an animal that looks like a bee has been given the insides of a spider, participants are more likely to judge the animal to be a spider when its telos changed to that of a spider compared to when its telos remained that of a bee. Moreover, when an animal that looks like a bee has acquired the telos of a spider, participants are more likely to judge the animal to be a spider when it has the insides of a spider as compared to when it has the insides of a bee. Nevertheless, "information about the insides seemed to be overall more informative of how the creature should be classified" (p. 10).

Notice first that if people do not categorize natural kinds solely by virtue of scientific essence, this alone would be enough to undermine the scientific essence view precisely because it takes scientific essence to be the sole criterion for category membership.

Notice second that someone who takes telos, or external features more generally, to be of main relevance to categorization can explain the found effect of insides just as well as a proponent of the scientific essence view appealing to diagnostic reasoning can explain the effect of telos. It can do so by assuming that participants make an inference from an animal having the insides of a category to that animal having the telos of that same category. Even though it is not part of the view that telos is caused by something inside the animal, when it comes to the most likely candidates for telos, it is a matter of common knowledge of biology. The fact that insides seem to have more of an effect than telos can subsequently easily be explained by appealing to the not implausible assumption that participants are more likely to infer that an animal has the telos of the category corresponding to the category of the insides when it

looks like the category of the insides compared to when it does not. It will additionally have to be assumed that participants take the conclusion of their inference to override what they are being told in the vignette. However, a scientific essence theorist appealing to diagnostic reasoning will have to make that assumption as well to explain away the results of R&N's second experiment.⁷

Further, J&Y's results can be said to count against the idea that participants reason diagnostically in the way Neufeld takes them to. In one of the assigned conditions, participants are told about an animal that looks like a bee, has the telos of a spider, and the insides of a bee. In this case, 74% of the participants judge that the animal in question still is a bee. But if the scientific essence view is true and participants make the inference assumed by Neufeld in R&N's second experiment, the animal should be judged to be of the new category because it has the telos of the new category, and having the telos of the new category should be diagnostic evidence of the scientific essence of the new category despite it being stated that it has the insides of the old category. This result, however, shows that participants do not do this. Hence, one might take this to show that the proponent of the scientific essence view cannot appeal to diagnostic reasoning in R&N's second experiment either. A defender of the scientific essence view might want to respond by saying that participants are more likely to infer insides from telos when the appearance of the animal corresponds to the category of the telos. But as seen in the second argument from Section 2, the proponent of the scientific essence view appealing to diagnostic reasoning must deny this assumption to explain the results of R&N's experiments.

In their second critical experiment, meant to address a third experiment by R&N, participants are told about a newborn bee that, after being hatched, is placed in a cage with spiders. Depending on the condition, participants are either told that the telos of the animal has changed to the purported telos of a spider, or that the insides of the animal have changed to that of a spider. A first important finding is that, contrary to the results of R&N, when the telos changes to that of a spider, 65% of the participants still classify the animal as a bee. A second important finding is that, similar to their previous experiment, the animal in question is more likely to be categorized as a spider when its insides change to that of a spider compared to when its telos changes to that of a spider.

Concerning their second finding, one might again take it to show that participants are more likely to infer that an animal has the telos of the category corresponding to the category of the insides when it looks like the category of the insides compared to when it does not.⁸

Concerning their first finding, it should be noted that J&Y's experiments are different from those of R&N in more respects than those thus far mentioned. In both J&Y's critical experiments, discrete response options are used, whereas R&N used a scale from 1 to 7, where 1 indicates that the animal in question is a bee and 7 indicates that it is a spider. Further, whereas in R&N's third experiment, participants are told about "skilled scientists" performing a "special experiment," J&Y's second critical experiment removed all talk of scientists and experiments. Arguably, the difference in results is a function of these changes. Notice, however, that removing the talk of scientists and experiments should, if arguments from diagnostic reasoning are to work, make diagnostic reasoning more likely. After all, it goes along with the removal of external manipulations and external manipulations are supposed to defeat the inference. As such, if these changes are to explain the results, and if that is to count in

favor of the scientific essence view, that should show a proponent of the scientific essence view that the issue is not with diagnostic reasoning.

5. Conclusion

The aim of this paper has not been to take a stance on the question of whether the scientific essence view or teleological essentialism is true. The aim, rather, is merely to point to the fact that the scientific essence view does indeed face a challenge. Perhaps there is an explanation available to the proponent of the scientific essence view that can explain the results of R&N and Haukioja et al.'s experiments. But an appeal to diagnostic reasoning, at least in the way it has been done thus far, is not it.

Acknowledgments

I am very grateful to James Hampton, Jussi Haukioja, and three anonymous referees, for very helpful feedback on earlier versions of this paper.

Notes

- 1 I will henceforth take for granted that participants take chemical composition to be responsible for the external features of substances like water.
- 2 R&N (2019) conduct five studies in total. Here, I will only be concerned with the first two, as they are the most important and addressing these will be enough to make my point. It is also noteworthy that subsequent studies on teleological essentialism seem to indicate that in addition to telos, appearance is of relevance to the categorization of natural kinds as well (Rose, Jaramillo, Nichols, & Horne, 2022; Zhang, She, Gerstenberg, & Rose, 2023). Nothing in what follows will hang on this.
- 3 I am thankful to a very helpful referee for the latter two points.
- 4 That category judgments of natural kinds are proportional to the amount of similarity in appearance and behavior an entity has to paradigmatic instances of a category is also one of Haukioja et al.'s findings.
- 5 In the second study, the mean is 1.97 in the “telos same” condition and 5 in the “telos changed” condition, where 1 indicates that the animal in question is definitely a bee and 7 indicates that the animal in question is definitely a spider. In the first study, the mean is 1.69 in the “telos same” condition and 5.58 in the “telos changed” condition. After performing a two-sided two-sample *t*-test by reversing the scale in one condition, both differences turned out to be statistically significant. $t(238) = 5.65, p = < .001$ for the second experiment and $t(227) = 3.48, p = < .001$ for the first experiment. Given that in both cases the animal in question is about equally more likely to be categorized as a bee than as a spider, the best explanation for this difference is that the purported bee telos is more indicative of being a bee than the purported spider telos is of being a spider.

- 6 Strictly speaking, for the appeal to diagnostic reasoning to work in R&N's second experiment, it will have to be assumed *either* that participants take the conclusion of their inference to override what they are being told in the vignette, *or* that participants do not take internal features to make for the scientific essence. Neufeld assumes that it is the latter because she thinks the scientific essence view is not committed to locating the scientific essence in the insides. However, it does not matter which disjunct is true because an appeal to arguments from diagnostic reasoning in the reverse Twin Earth cases would have to rely on the same disjunctive assumption.
- 7 Notice that J&Y assume that the scientific essence is located in the insides, for without that assumption, the effect of insides cannot count in favor of the scientific essence view. As such, they are also committed to the first disjunct of the disjunctive assumption in fn. 6.
- 8 Although there is no statement that makes explicit what kind of insides the animal has in one of the conditions, they are implied to be the insides of the original category by not stating that the insides have changed.

References

- Braisby, N., Franks, B., & Hampton, J. A. (1996). Essentialism, word use, and concepts. *Cognition*, *59*, 247–274.
- Devitt, M., & Porter, B. C. (2021). Testing the reference of biological kind terms. *Cognitive Science*, *45*(5), e12979.
- Fernbach, P. M., Macris, D. M., & Sobel, D. M. (2012). Which one made it go? The emergence of diagnostic reasoning in preschoolers. *Cognitive Development*, *27*(1), 39–53.
- Gelman, S. (2003). *The essential child*. New York: Oxford University Press.
- Hampton, J. A., Estes, Z., & Simmons, S. G. (2007). Metamorphosis: Essence, appearance and behavior in the categorization of natural kinds. *Memory & Cognition*, *35*, 1785–1800.
- Haukioja, J., Nyquist, M., & Jylkkä, J. (2021). Reports from Twin Earth: Both deep structure and appearance determine the reference of natural kind terms. *Mind & Language*, *36*(3), 377–403.
- Joo, S., & Yousif, S. R. (2022). Are we teleologically essentialist? *Cognitive Science*, *46*(11), e13202.
- Keil, F. C. (1989). *Concepts, kinds and cognitive development*. Cambridge, MA: MIT Press.
- Kripke, S. A. (1980). *Naming and necessity*. Oxford: Basil Blackwell.
- Machery, E., Olivola, C. Y., Cheon, H., Kurniawan, I. T., Mauro, C., Struchiner, N., & Susianto, H. (2023). Is identity essentialism a fundamental feature of human cognition? *Cognitive Science*, *47*(5), e13292.
- Malt. (1994). Water is not H₂O. *Cognitive Psychology*, *27*(1), 41–70.
- Neufeld, E. (2021). Against teleological essentialism. *Cognitive Science*, *45*(4), e12961.
- Putnam, H. (1975). The meaning of 'meaning'. In H. Putnam (Ed.), *Philosophical papers* (pp. 215–271). Cambridge: Cambridge University Press.
- Rose, D., & Nichols, S. (2019). Teleological essentialism. *Cognitive Science*, *43*(4), e12725.
- Rose, D., Jaramillo, S., Nichols, S., & Horne, Z. (2022). Teleological essentialism across development. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *44*. Retrieved from <https://escholarship.org/uc/item/0z45p9zs>
- Sobel, D. M., Yoachim, C. M., Gopnik, A., Meltzoff, A. N., & Blumenthal, E. J. (2007). The blicket within: Preschoolers' inferences about insides and causes. *Journal of Cognition and Development*, *8*(2), 159–182.
- Tobia, K., Newman, G. E., & Knobe, J. (2020). Water is and is not H₂O. *Mind & Language*, *35*(2), 183–208.
- Zhang, S., She, J. S., Gerstenberg, T., & Rose, D. (2023). You are what you're for: Essentialist categorization in large language models. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *45*. Retrieved from <https://escholarship.org/uc/item/3996v30z>