

広島大学学術情報リポジトリ  
Hiroshima University Institutional Repository

Title	Inferentialist ethics for analyzing mathematical interactions in the age of artificial intelligence
Author(s)	Uegatani, Yusuke; Otani, Hiroki
Citation	秋期研究大会発表集録 , 52 : 121 - 124
Issue Date	2019-11-16
DOI	
Self DOI	
URL	<a href="http://ir.lib.hiroshima-u.ac.jp/00048312">http://ir.lib.hiroshima-u.ac.jp/00048312</a>
Right	Copyright (c) 2019 日本数学教育学会
Relation	



# Inferentialist ethics for analyzing mathematical interactions in the age of artificial intelligence

Yusuke UEGATANI  
Hiroshima University High School, Fukuyama

Hiroki OTANI  
Japan Women's University

## 1. Introduction

Inferentialism, which is proposed as a pure philosophy by R. Brandom (2000), has increased its influence on educational research, particularly mathematics education. It sheds light on the rationality of everyday concepts (Derry, 2008) and acknowledges the importance of what is already implied in traditional educational discourses (Su & Bellmann, 2018). It questions the classic view of an implicit assumption about the orders of conceptual development in mathematics learning (Bakker & Derry, 2011) and captures human concept development in terms of holistic concept uses in social practices (Hußmann, Schacht, & Schindler, 2018; Uegatani & Otani, in press).

One of the most impactful proposals from an inferentialist perspective is integrating the famous two metaphors for learning, acquisition and participation (Sfard, 1998), into one metaphor entitled mastering (Taylor, Noorloos, & Bakker, 2017). It characterizes learning as the duality of acquiring socially acknowledged normative concepts and the participation in rational engagements with these concepts. It may even become a new foundation of educational research. However, the authors of the present paper would like to argue that all of the three metaphors for learning fail to capture the nature of human mathematical learning, especially in comparison with machine learning. An artificial intelligence (AI), which has a high-performance processing unit, a highly sensitive sensor, a highly functional body, and a sufficient ability to build a society as a network between other AIs, might acquire many pieces of knowledge as data and gradually improve practical and social behaviors in a statistical way. Since the metaphor of mastering is an integrated idea of the famous two metaphors of acquisition and participation, it essentially cannot surpass its inherent limitation: the involvement of AI and its capacity to master a given subject.

This is a serious claim especially in the age of AI. If everything learnable by humans is also learnable by AI, then there might be no need for humans to learn anything. In order to discuss and make explicit what humans should or can learn especially regarding mathematics, we need to discover ways to understand the contrasts between the unique characteristics of human learning and machine learning.

The purpose of this study is to propose an ethical perspective compatible with inferentialism in mathematics education research. The reason for focusing on intentionality is that although it is one of the core elements in inferentialism and important for characterizing what is distinctively human in inferentialist discourses, its role in human learning is not explicitly discussed in Taylor, Noorloos, & Bakker's (2017) mastering metaphor proposal.

## 2. Key ideas in inferentialism

Inferentialism is linguistic pragmatism, which considers the meaning of a word as its inferential role rather than as its representational role. Inferentialists in education focus on how a student uses a mathematical term (a mathematical concept) for reasoning rather than what the term (the concept) should represent in school mathematics (Bakker & Derry, 2011). Since inferentialists have an interest in the actual ways of using concepts, they apply holism to human

concept use. In addition, they also think that when one concept is used, many other concepts are involved simultaneously and implicitly (Brandom, 2000, p. 15). This applies especially to students who are not completely conscious of which concepts, and how many, they use at a given moment. For this reason, inferentialists think that “[t]he process of explicitation is to be the process of applying concepts: conceptualizing some subject matter” (Brandom, 2000, p. 8). Therefore, in educational research regarding a student’s conceptualizations, inferentialists gather data on *what a student makes explicit*. Inferentialists do not believe that students can express a mathematical concept because they understand it; rather, they think that students understand a mathematical concept because they can express it (cf. Uegatani & Otani, in press).

Inferentialism is also a philosophy aimed at characterizing what is distinctively human. One of the key elements of this characterization is the idea of *intentionality* as well as the ideas of *aboutness* and *de re ascriptions*. “The prime explicitly representational locution of natural language is *de re ascriptions of propositional attitudes*. It is their use in these locutions that make the word ‘of’ and ‘about’ express the intentional directions of thought and talk” (Brandom, 2000, p. 169). Let  $\phi(x)$  be a sentence including the term  $x$ ,  $S$  be a subject, and  $t$  be a term. Then, the *de re* form of the fact that  $S$  believes that  $\phi(t)$  is as follows:  $S$  believes *of*  $t$  that  $\phi(it)$  (Brandom, 2000, p. 172). For example, when Tom believes that two is the smallest prime number, we can rewrite this fact from the observer standpoint: Tom believes *of* the smallest prime number that it is two; or, Tom believes *of* two that it is the smallest prime number. Thus, *de re* ascription attributes *what a person talks about* to *what he/she makes explicit* from the observer standpoint. The aboutness of what a person makes explicit regards the direction he/she intends to take.

### 3. Intentionality and AIs

#### (1) Ethics and human intentionality

Although we can think that an AI intends to talk about something, it actually does not in principle. Since all AI behavior is determined in a computable manner, its behavior deterministically depends on how the environment and its inner state change. Thus, its behavior is not intended by itself as it is nothing more than an output or a reaction to an input. This is a fundamental limitation common to any artifact. In other words, we can only discuss whether or not the AI behaves in accordance with the intention of the AI designers. If the designers do not hold any expectations on the AIs functioning, then it never learns as we expect, and something becomes learnable by humans, but not the AI. Therefore, in order to understand any human behavior, one of the important distinctions between humans and artifacts is intentionality.

Considering this distinction, some readers may raise the following question: How can we capture what a student intends to talk about? As radical constructivists like von Glasersfeld (1990) correctly recognized that humans are informationally closed. We cannot directly access another’s intentionality.

This problem has caused a delicate tension between cognitive and social perspectives in mathematics education. For example, Ernest (2012) denounced radical constructivists as unethical because they treat students as mechanical thinkers or they disregard students’ intentionality. However, this critique does not imply that there is a way of capturing the correct human intentionality. If such a way exists, it might also be applicable to capturing the apparent but non-existent intentionality of AIs.

We believe that we can never correctly know what an observed person intends to talk about. Instead, we make two assumptions: 1) *every human or human-like entity has intentionality*; and 2) applying the first assumption to ourselves, *we as researchers also have intentionality*. Due to these assumptions, we avoid an ontological problem regarding the existence of true intentionality rather than solve it. In the following two sections, we explain the ethical validity of these assumptions in relation to inferentialism.

## (2) The ethical validity of the first assumption

The first assumption is ethically valid because even an AI, which behaves like humans, should be considered to have intentionality. M. Coeckelbergh (2018), a contemporary philosopher of technology who is known for his application of Wittgenstein's philosophy to the ethics of robotics and AI, argues that the moral standing of AIs depends on how we talk about it. For example, if a human-like robot is abused, people often feel pity. From an inferentialist holistic view on human concept use, this means that the degree of similarity between the concepts used for describing an AI and a human determines the degree to which the AI is treated like a human. Also, Ernest (2012) and Coeckelbergh (2018) both claim that ethics should precede ontology from the Levinasian perspective. The ontological fact that an AI is actually just a non-human object should not influence our initial ethical attitude toward it.

Furthermore, assuming the existence of intentionality behind an observed human behavior (or human-like behavior) is an ethical requirement for researchers involving human subjects. On the other hand, assuming that we can correctly determine or precisely control what a person intends to talk about is unethical. Every human (or human-like entity) has a freedom to think without violation from others. We as educational researchers are thus required to have a careful attitude when assuming that a true intentionality of a person exists. However, this true intentionality is unknowable, thus, what we can and should do is build a hypothetical intentionality of the person. Since it is ethically valid for us to create a non-existent hypothetical intentionality of AIs, it is also ethically valid to construct a hypothetical intentionality of the person even if it does not match true intentionality. We should always continue to update temporal hypothetical intentionality in order to treat the observed persons as humans. In other words, we should continue to revise our viable model for intentionality of a specific person in the radical constructivist epistemological sense (cf. Steffe & Thompson, 2000).

## (2) The ethical validity of the second assumption

The second assumption is ethically valid because we as researchers are also humans. Our research practice is a kind of purposeful human social activity. It cannot exist independent from our intentionality, i.e., neglecting what we intend to talk about. In addition, it is highly important, in this paper, that the validity of hypothetical intentionality depends on our intentionality, though this claim is a little bit radical from the classic natural scientific point of view. Since our research practice is not computation, like the behaviors of AIs, we refrain from obtaining our result in a deterministic manner. Mechanically restricting our research practices against our own intentionality is also unethical. This is a consequence of the idea that depending on our purpose, we as educational researchers need to ascribe *de re* to the claims of an observed student.

For example, suppose that when a student is asked if a given curved graph is a graph of  $y=2x+3$  she claims that "the graph of linear functions should be a straight line." Then, although she may be conscious of her intent to talk about the graph of  $y=2x+3$ , we as the observers can make different valid interpretations of her claims. For instance, 1) she believes of a linear function that it should be a straight line; 2) she believes of a curved graph that it should not be a graph of a linear function; or 3) she believes of a straight line that it is a necessary condition for the graph of a linear function. Independent from what she actually intended to talk about, we can, to some extent, arbitrarily ascribe *what she talked about* to *what she actually said*.

Of course, this ascription *de re* is not completely arbitrary. Her above claim cannot be about, for example, her mother. However, since we as inferentialists emphasize not formal inferences, but material inferences, we cannot define what a good ascription *de re* is in general. In addition, we only focus on what is made explicit. We cannot estimate the intentionality of a person before observing any behavior and we should always consider what the observed student said explicitly.

#### 4. Conclusion

To summarize this paper, we argue the necessity of two assumptions that are also ethical requirements for us: 1) *every human or even human-like entity has intentionality*; and 2) applying the first assumption to ourselves, *we as researchers also have intentionality*. In addition, avoiding the ontological problem of true intentionality, we discussed our epistemological ways of building hypothetical intentionality. We may, to some extent, arbitrarily ascribe *de re* to how the observed students explicitly behaved, which is independent from their true intentionality. Note that there are two limitations to our proposal. First, the present paper argues the possibility of inferentialism only from a theoretical point of view. We must actualize this possibility through more empirical research. Second, we did not solve the problem of learning metaphors, though it was our starting point. We must tackle this issue through the findings of future empirical research.

#### Acknowledgements

This work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number 18K13162 and 18H05751. We would like to thank Editage ([www.editage.jp](http://www.editage.jp)) for English language editing.

#### References

- Bakker, A., & Derry, J. (2011). Lessons from Inferentialism for Statistics Education. *Mathematical Thinking and Learning*, 13(1–2), 5–26. <https://doi.org/10.1080/10986065.2011.538293>
- Brandom, R. (2000). *Articulating Reasons*. Harvard University Press.
- Coeckelbergh, M. (2018). Why Care About Robots? Empathy, Moral Standing, and the Language of Suffering. *Kairos. Journal of Philosophy & Science*, 20(1), 141–158. <https://doi.org/10.2478/kjps-2018-0007>
- Derry, J. (2008). Abstract rationality in education: From Vygotsky to Brandom. *Studies in Philosophy and Education*, 27(1), 49–62. <https://doi.org/10.1007/s11217-007-9047-1>
- Ernest, P. (2012). What is our first philosophy in mathematics education? *For the Learning of Mathematics*, 32(3), 8–14.
- Hußmann, S., Schacht, F., & Schindler, M. (2018). Tracing conceptual development in mathematics: Epistemology of webs of reasons. *Mathematics Education Research Journal*, 1–17. <https://doi.org/10.1007/s13394-018-0245-7>
- Sfard, A. (1998). On Two Metaphors for Learning and the Dangers of Choosing Just One. *Educational Researcher*, 27(2), 4–13. <https://doi.org/10.3102/0013189X027002004>
- Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In R. Lesh & A. E. Kelly (Eds.), *Handbook of Research Design in Mathematics and Science Education* (pp. 267–306). Hillsdale, NJ: Erlbaum.
- Su, H., & Bellmann, J. (2018). Inferentialism at Work: The Significance of Social Epistemology in Theorising Education. *Journal of Philosophy of Education*, 52(2), 230–245.
- Taylor, S. D., Noorloos, R., & Bakker, A. (2017). Mastering as an Inferentialist Alternative to the Acquisition and Participation Metaphors for Learning. *Journal of Philosophy of Education*, 51(4), 769–784. <https://doi.org/10.1111/1467-9752.12264>
- Uegatani, Y., & Otani, H. (in press). A new ontology of reasons for inferentialism: Redefining the notion of conceptualization and proposing an observer effect on assessment. *Mathematics Education Research Journal*. <https://doi.org/10.1007/s13394-019-00289-8>
- von Glasersfeld, E. (1990). Distinguishing the Observer: An Attempt at Interpreting Maturana. In V. Riegas & C. Vetter (Eds.), *Zur Biologie der Kognition* (pp. 281–295).