

Comment on “Algorithms and Authenticity” by Arthur S. Jago

Glenn R. Carroll (Stanford University) and Kieran O’Connor (University of Virginia)

Arthur S. Jago’s article on algorithms and authenticity documents a compelling—and even startling—new set of empirical findings. He finds that when encountering products or services delivered by artificial intelligence (AI) algorithms, individuals often react and interpret the interaction through the lens of authenticity. Compared to humans, an algorithm’s perceived authenticity is lower, although most of this occurs because of low perceived moral authenticity. To restate casually, individuals are fine with algorithms performing analytical functions but less happy when they are doing things that people believe should take into account the values and moral principles inherent in sentient beings.

The evidence Jago presents makes his study compelling. The domain of application makes it startling. While many have noted the increasing prevalence of authenticity as a taste filter in modern society, few have imagined that the reach of authenticity would be so broad—most current research applications are to markets of personal consumption and culture.

Jago’s study adds to a growing body of theory and research showing the varied—but systematic—meanings of the linguistic term *authenticity* used by consumers and analysts. Jago’s findings on type and moral authenticity rest comfortably aside our findings on the same two concepts in the context of food items, handbags, and restaurants (O’Connor and colleagues, 2017). Moreover, Lehman and colleagues (2017) develop a validated scale for type and moral authenticity, which might prove more useful than Jago’s self-admitted ad-hoc two-item scales. Newman and Smith (2016) describe a

variety of other ways to discriminate among the meanings of authenticity currently invoked by consumers.

When Jago's study is considered side by side with many other studies of authenticity, it demonstrates (again) that at an abstract level, context matters hugely for interpretation. Clearly, authenticity interpretations vary by market, social setting, product and service, perhaps in interactive ways. We suspect strongly that societal context, even national culture, matters too, and we look forward to seeing future studies investigating the issue. Ultimately, the scientific goal would be a theory that makes predictions based on these many contingencies, although it may be a way off.

The theoretical challenges are complex. Jago shows that faced with an AI producer, individuals use implicit ideas (lay theory) about authenticity to interpret the situation and use type authenticity to value the producer positively. But the relationship between AI and type authenticity is clearly not symmetric. That is, when an individual is faced with a type authentic cultural entity, it does not seem likely to imply that he/she will regard AI as a necessarily acceptable producer. Consider your favorite category-conforming Thai restaurant: it is type authentic. While the recipes used might be algorithms of a sort, we doubt that patrons would be comfortable with food made and served by robots, or that they would even be as comfortable with comparable dishes made or served by human chefs, as his study implies.

Jago's study also shows the framing or story matters a lot about how individuals interpret the meaning of authenticity. Many product, service, or place stories are often implicit and hard for consumers to access. Jago shows that making them explicit might help managers see and manage consumer reaction. We agree and think this line of

thinking can be extended. Jago focuses mainly on comparing algorithms and humans, which generates starkly different interpretations. But we also wonder about what kinds of authenticity interpretations, if any, are made when the alternatives are all algorithms, just different algorithms. An example of current prominence would be the models used to analyze and predict hurricanes. As currently presented by the news, these are different sterile AI models, but we suspect that if the stories of their development and their assumptions were brought to life, some would seem more authentic than others. Even anthropomorphizing actual hurricanes seems to affect responses to them, such as when individuals tend to prepare less for, and suffer more from, hurricanes with female compared to male names (Jung and colleagues, 2014). What Jago's work uncovers about authenticity judgments might be less about algorithms per se, and more about attributions driven by what is human (or even human-seeming) vs. non-human.

We end by making an observation about the possible implication of Jago's study for evaluating AI. For decades, a standard for evaluating AI has been the famous Turing test, which relies mainly on natural language, reason, knowledge and learning. Jago's study suggests that interpretation also depends on reason using values and moral principles, that is, judgment. If so, then the agenda for AI becomes much bigger and more challenging.

Jung, K., Shavitt, S., Viswanathan, M., & Hilbe, J. M. 2014. Female hurricanes are deadlier than male hurricanes. *Proceedings of the National Academy of Sciences*, 111(24): 8782–7.

Lehman, D. W., O'Connor, K., & Carroll, G. R. 2017. Acting on authenticity: Audience responses to varying meanings. Paper presented at the annual meetings of the Academy of Management, Atlanta.

Newman, G. E., & Smith, R. K. 2016. Kinds of authenticity. *Philosophy Compass*, 11(10): 609–618.

O'Connor, K., Carroll, G. R., & Kovács, B. 2017. Disambiguating authenticity: Interpretations of value and appeal. *PLOS ONE*, 12(6): e0179187.