

**Individual and Cross-Cultural Differences in Semantic Intuitions:
New Experimental Findings***

James R. Beebe (University at Buffalo) and Ryan J. Undercoffer (Syracuse University)

[Under Review]

In 2004 Edouard Machery, Ron Mallon, Shaun Nichols, and Stephen Stich published what has become one of the most widely discussed papers in experimental philosophy, in which they reported that East Asian and Western participants had different intuitions about the semantic reference of proper names. A flurry of criticisms of their work has emerged, and although various replications have been performed, many critics remain unconvinced. We review the current debate over Machery et al.'s (2004) results and take note of which objections to their work have been satisfactorily answered and which ones still need to be addressed. We then report the results of studies that reveal significant cross-cultural and intra-cultural differences in semantic intuitions when we control for variables that critics allege have had a potentially distorting effect on Machery et al.'s findings. These variables include the epistemic perspective from which participants are supposed to understand the research materials, unintended anchoring effects of those materials, and pragmatic factors involved in the interpretation of speech acts within them. Our results confirm the robustness of the cross-cultural differences observed by Machery et al. and thereby strengthen the philosophical challenge they pose.

keywords: semantics, philosophy of language, cross-cultural, experimental philosophy

* We would like to thank...

1. The Original Findings

Machery et al. (2004) investigated the question of whether East Asian and Western individuals had intuitions about proper names that lined up with either descriptivist or causal-historical theories of reference. According to descriptivism, speakers associate a description with each proper name, and the name refers to an object if that object uniquely or best satisfies the description. Importantly, if the description is not satisfied by any object or if several objects equally satisfy the description, the name fails to refer. According to the causal-historical view, a name may be introduced into a linguistic community with a description, but after this initial “baptism” the description plays no subsequent role in determining its referent. Successive uses of the term refer to the same object if those uses are connected to this first use via an appropriate causal chain. Notably, an object may fail to satisfy any description a language user may associate with the name and yet succeed in being its referent.

Machery et al. (2004) presented 31 Rutgers University undergraduates of Western European ancestry and 40 undergraduates from the University of Hong Kong whose native language was Chinese the following version of Kripke’s (1972/1980, pp. 83-92) Gödel/Schmidt case:

Gödel. Suppose that John has learned in college that Gödel is the man who proved an important mathematical theorem, called the incompleteness of arithmetic. John is quite good at mathematics and he can give an accurate statement of the incompleteness theorem, which he attributes to Gödel as the discoverer. But this is the only thing that he has heard about Gödel. Now suppose that Gödel was not the author of this theorem. A man called ‘Schmidt,’ whose body was found in Vienna under mysterious circumstances many years ago, actually did the work in question. His friend Gödel somehow got hold of the manuscript and claimed credit for the work, which was thereafter attributed to Gödel. Thus, he has been known as the man who proved the incompleteness of arithmetic. Most people who have heard the name ‘Gödel’ are

like John; the claim that Gödel discovered the incompleteness theorem is the only thing they have ever heard about Gödel. When John uses the name ‘Gödel,’ is he talking about:

(A) the person who really discovered the incompleteness of arithmetic? or

(B) the person who got hold of the manuscript and claimed credit for the work?

Machery et al. also gave each participant a structurally analogous case that involved a Chinese name and context. The full text of this case (Gödel2) is found in the Appendix below. In what we will refer to as ‘the two Gödel cases,’ there is a description that contemporary speakers associate with a name but that is allegedly satisfied by someone other than the original bearer of the name. As a result, descriptivism seems to require that (A) is the correct answer in both cases. According to the causal-historical view, however, (B) is correct because contemporary speakers are historically related to the original bearer of the name in question.

Machery et al. (2004) also gave each participant the following version of Kripke’s (1972/1980, pp. 66-67) Jonah case:

Jonah1. In high-school, German students learn that Attila founded Germany in the second century A.D. They are taught that Attila was the king of a nomadic tribe that migrated from the east to settle in what would become Germany. Germans also believe that Attila was a merciless warrior and leader who expelled the Romans from Germany, and that after his victory against the Romans, Attila organized a large and prosperous kingdom.

Now suppose that none of this is true. No merciless warrior expelled the Romans from Germany, and Germany was not founded by a single individual. Actually, the facts are the following. In the fourth century A.D., a nobleman of low rank, called ‘Raditra,’ ruled a small and peaceful area in what today is Poland, several hundred miles from Germany. Raditra was a wise and gentle man who managed to preserve the peace in the small land he was ruling. For this reason, he quickly became the main character of many stories and legends. These stories

were passed on from one generation of peasants to the next. But often when the story was passed on the peasants would embellish it, adding imaginary details and dropping some true facts to make the story more exciting. From a peaceful nobleman of low rank, Raditra was gradually transformed into a warrior fighting for his land. When the legend reached Germany, it told of a merciless warrior who was victorious against the Romans. By the eighth century A.D., the story told of an Eastern king who expelled the Romans and founded Germany. By that time, not a single true fact remained in the story.

Meanwhile, as the story was told and retold, the name ‘Raditra’ was slowly altered: it was successively replaced by ‘Aditra,’ then by ‘Arritrak’ in the sixth century, by ‘Arrita’ and ‘Arrila’ in the seventh and finally by ‘Attila.’ The story about the glorious life of Attila was written down in the eighth century by a scrupulous Catholic monk, from whom all our beliefs are derived. Of course, Germans know nothing about these real events. They believe a story about a merciless Eastern king who expelled the Romans and founded Germany.

When a contemporary German high-school student says “Attila was the king who drove the Romans from Germany,” is he actually talking about the wise and gentle nobleman, Raditra, who is the original source of the Attila legend, or is he talking about a fictional person, someone who does not really exist?

(A) He is talking about Raditra.

(B) He is talking about a fictional person who does not really exist.

Again, a structurally similar case featuring a Chinese name and context was also employed (cf. Jonah² in the Appendix). According to descriptivism, the correct answer in both Jonah cases is (B) because there is no one who satisfies the description associated with the names in question. On the causal-historical view, however, satisfying the description is not necessary for the name to denote anyone, and so (A) is the correct answer. Following Machery et al. we call participant responses that are consistent

with the causal-historical view and inconsistent with descriptivism ‘Kripkean responses’ and responses that are consistent with descriptivism but not the causal-historical theory ‘descriptivist responses.’

Machery et al. (2004) computed a Kripkean intuition score for each participant by giving them a ‘1’ for each Kripkean intuition and a ‘0’ for each descriptivist intuition they expressed. Mean Kripkean intuition scores are reported in Table 1.

	Westerners	East Asians
Gödel cases	1.13	.63
Jonah cases	1.23	1.32

Table 1. Mean Kripkean intuition scores for Western and East Asian participants in Machery et al.’s (2004) cross-cultural study.

The difference between the mean scores for Western and East Asian participants in relation to the Gödel cases was statistically significant, but there were no significant differences between Western and East Asian participants concerning the Jonah cases. More recently, Machery (2012, 40) reports the following additional details about how East Asian and Western participants responded to the two Gödel cases:

	Westerners	East Asians
Gödel1	58%	29%
Gödel2	55%	32%

Table 2. Percentages of Kripkean intuitions given by Western and East Asian participants in response to the two Gödel cases in Machery et al. (2004).

On the basis of these results Machery et al. (2004, B1) conclude:

The experiment indicated that, for certain central cases, Westerners are more likely than East Asians to report intuitions that are consistent with the causal-historical view. These results constitute prima facie evidence that semantic intuitions vary from culture to culture, and the paper argues that this fact raises questions about the nature of the philosophical enterprise of developing a theory of reference.

The results in the Gödel cases have been replicated by Machery et al. (2009) and Machery, Deutsch, and Sytsma (forthcoming). Machery et al. (2010) also replicated the Gödel results using Chinese translations and Chinese participants. In addition, Livengood, Sytsma, Sato, and Mineki (unpublished data) have replicated Machery et al.'s (2004) Gödel results with Japanese participants, finding that Japanese tend to have descriptivist intuitions.

In what follows we survey the most prominent objections that have been raised to the work of Machery and his collaborators, beginning with those objections that we believe have been satisfactorily answered or at least do not pose a serious threat (Section 2). We then explain some issues that remain unresolved and the experiments we designed to address them. One weakness of the replications that have been performed is that they have focused on only one Gödel case at a time and have neglected the Jonah cases. Thus, we began our investigation of cross-cultural semantic intuitions with an attempted replication of all of Machery et al.'s (2004) results. Although we initially failed to find cross-cultural differences in participants' responses to the Jonah cases (Section 3), we eventually did find such differences when we modified the probe questions that participants were asked (Section 7). We also report measures of the extent of intra-cultural variation in each of our studies, one of which reveals significant correlations between certain personality traits and participants' semantic intuitions (Sections 4 and 8).

One of the most significant objections that has been raised against the work of Machery et al. is that it is not clear whether participants should understand the Gödel cases from the limited epistemic perspective of the central protagonist or from the omniscient narrator's perspective (Sytsma and Livengood, 2011). In Section 5, we show that cross-cultural differences remain when the proper perspective is clarified. An additional concern about Machery et al.'s Gödel cases is that they may unwittingly bias participants against what are supposed to be the descriptivist responses by providing misleading anchors at the beginning of the vignettes. In Section 6, we report the results of using modified vignettes that control for this possibility and show that cross-cultural differences in semantic intuitions remain. By providing empirical evidence that many of the worries expressed by critics do not undermine the original results of Machery et al., our results demonstrate the robustness of the original findings and strengthen the philosophical challenge they pose.

2. Answered or Non-Threatening Objections

The work of Machery et al. has received substantial criticism from a variety of different angles. Some object to Machery et al.'s (2004, B3) claim that “[t]here is widespread agreement among philosophers on the methodology for developing an adequate theory of reference,” which they describe as follows:

The project is to construct theories of reference that are consistent with our intuitions about the correct application of terms in fictional (and non-fictional) situations. Indeed, Kripke's masterstroke was to propose some cases that elicited widely shared intuitions that were inconsistent with traditional descriptivist theories (ibid.).

Mallon et al. (2009, p. 338) make similar claims:

We propose that to find the correct theory of reference, philosophers of language are committed to using what is sometimes called ‘the method of cases’:... The correct theory of reference for a

class of terms *T* is the theory which is best supported by the intuitions competent users of *T* have about the reference of members of *T* across actual and possible cases.

Michael Devitt (2011, pp. 419, 421) grants that theories of reference are standardly evaluated by consulting intuitions but argues that Machery et al. have “exaggerate[d] the role of intuitions about the *reference* of terms in *hypothetical* situations” and that “[t]he further one gets away from actual error toward hypothetical cases that may seem a bit fanciful, the less confident we should be about our own intuitions and... *especially, those of the folk.*” Devitt (2011, p. 419) also argues that Machery et al. neglect the fact that many of the key intuitions Kripke tries to elicit “are not semantic ones about reference but metaphysical ones about *modal properties.*”

In a similar vein, Jonathan Ichikawa, Ishani Maitra, and Brian Weatherson (2012) argue that the results of Machery et al.’s experiment do less damage to Kripke’s overall critique of descriptivism than the experimentalists appear to think.¹ Likewise, Max Deutsch (2009, pp. 445, 448) contends that “nothing in Kripke’s famous argument against the *descriptivist theory of reference* for proper names hinges on assuming anything about peoples’ intuitions” and that “the predictions of a theory of reference concern terms and their referents, not competent speakers and their intuitions.” It is our opinion that the work of Machery et al. remains significant, regardless of the particular details of Kripke’s case against descriptivism and regardless of what the proper methodology for semantics and the philosophy of language may be. If the Kripkean perspective is supposed to account for the reference of proper names, as those names are employed in ordinary linguistic practice, we think it is an interesting question how this perspective can accommodate cross-cultural and intra-cultural variation in semantic intuitions. Consequently, we will not concern ourselves in what follow with

¹ Cf. Machery et al. (2013) for a response to Ichikawa et al.’s arguments.

objections to Machery et al. that depend upon the finer points of interpretation of *Naming and Necessity*.²

Genoveva Martí (2009, p. 44, 45) has argued that by asking participants about the semantic properties of a mentioned (rather than a used) name, Machery et al. were actually testing “people’s intuitions about *theories* of reference, not about the *use* of names,” whereas the central question was supposed to be “whether the East Asian participants *use* names descriptively.” Machery et al. (2009, 689) offer the following explanation of the relevant distinction:

In Kripke’s Gödel case (1972/1980, 83-84), the judgement that the proper name ‘Gödel’ refers to Gödel and not to Schmidt is a metalinguistic intuition, since it is about the reference of the proper name ‘Gödel’; by contrast, the judgement that in this case Gödel should not have claimed credit for the incompleteness theorem is a linguistic intuition.

Machery et al. (2009) responded to this criticism by collecting further empirical data on the Gödel case that showed that the metalinguistic intuitions of participants from India, France, and Mongolia did not differ significantly from their linguistic intuitions. Machery and Olivola (unpublished data) uncovered a similar finding with American participants. The question Machery et al. (2009) designed to elicit linguistic intuitions was the following:

Having read the above story and accepting that it is true, when Ivy says, ‘Tsu Ch’ung Chih was a great astronomer,’ do you think that her claim is: (A) true or (B) false?

The metalinguistic question that Machery et al. (2009) used was the following:

Having read the above story and accepting that it is true, when Ivy uses the name ‘Tsu Ch’ung Chih,’ who do you think she is actually talking about:

(A) the person who (unbeknownst to Ivy) really determined the solstice times?

or

² Cf. Machery et al. (2013) for further discussion of these issues.

(B) the person who is widely believed to have discovered the solstice times, but actually stole this discovery and claimed credit for it?

Their results appear in Table 3.

	U.S.A.	India	France	Mongolia
Kripkean linguistic intuitions	64%	67%	56%	66%
Kripkean metalinguistic intuitions	67%	58%	48%	56%

Table 3. Percentages of Kripkean linguistic and metalinguistic intuitions in Machery et al.'s (2009) and Machery and Olivola (unpublished) cross-cultural studies.³

Machery et al. (2009, p. 693) note that they found substantial within-culture variation in intuitions about reference in all three countries and argue that “this within-culture variation poses exactly the same philosophical challenge as cross-cultural variation.”⁴ Not only did Machery et al. (2009) find no statistically significant difference between the linguistic and metalinguistic intuitions of participants within each culture, they also found no significant differences between the responses of Indian, French, and Mongolian cultures. The semantic intuitions of American participants, however, did differ from those of some of the other sets of participants:

Consistent with previous work, Americans are more likely to have causal-historical intuitions about the reference of proper names than people from Mongolia, India, and France when asked about what a proper name refers to, and they were more likely than French people to have causal-historical intuitions when asked about the truth-value of a sentence containing proper names. (Machery and Stich 2012, p. 502)

³ Although Machery et al. (2009, p. 693) claim that “regardless of the country, about two thirds of our participants tended to have Kripkean intuitions,” it is not clear how this description applies to the French participants whose mean responses were 48% and 56%.

⁴ Machery et al. (2004) make a similar point.

One noteworthy feature of the French responses is that—unlike the (western) American responses in Machery et al. (2004) that showed a marked preference for the Kripkean view—the (western) French responses in the present experiment are split roughly in half, thereby displaying significant intra-cultural variation.⁵ Although Martí (forthcoming) remains unconvinced, the consistency that Machery and his collaborators have demonstrated between linguistic and metalinguistic intuitions in their samples suggests that there is not a fatal flaw in this aspect of their original experiment.

On another front, Barry Lam (2010, 321-322) has criticized Machery et al. (2004) for the fact that their research materials were written entirely in English:

The most straightforward way of questioning this presumption [regarding the universality of semantics intuitions] would be to test Kripke's theory of reference as a theory of Cantonese (or Mohawk or Swahili) names by asking native Cantonese (or Mohawk or Swahili) speakers in their native-languages about the referents of Cantonese (or Mohawk or Swahili) names when used by Cantonese (or Mohawk or Swahili) speakers in hypothetical cases.... Cross-cultural differences resulting from a study of intuitions wholly in English admit of alternative linguistic explanations, as opposed to intuitional or psychological explanations. For instance, rather than genuine differences in intuitions, the differences in responses to the questions might result from differences in native-language and second-language competency.

In response to Lam's criticisms, Machery et al. (2010) translated Gödel1 into Chinese and replicated their original finding. Chinese participants were still more inclined to give a descriptivist answer than a causal-historical one and native English speakers were still more likely to give a causal-historical answer. Machery et al. (2010) also gave Chinese participants both English and Chinese translations of one of Lam's vignettes and found that their answers did not differ when different languages were used,

⁵ Cf. Machery (2011) and Martí (forthcoming) for further discussion of these issues.

contrary to Lam’s initial suggestion.⁶ Sytsma, Livengood, Sato, and Mineki (unpublished data) have also found that Japanese participants were likely to have descriptivist intuitions about Gödel1.

One of the most common responses to the work of Machery et al. (2004)—and indeed to much experimental philosophy—has been the Expertise Defense, whereby critics charge that empirically investigating the intuitions of untrained participants is not relevant to the true task of philosophy and poses no challenge to philosophical methodology as we know it. Devitt (2011, p. 424), for example, writes, “No biologist thinks that the systematization of our ordinary intuitions about living things is the way to do biology. There is no good reason to think that semantics is different.” Ludwig (2007) and Deutsch (2009) express similar sentiments in response to Machery et al.’s (2004) work.

In a fascinating follow-up experiment, Machery (2012) recruited linguists and philosophers of language from around the world to register their intuitive response to Gödel2. Of the 272 scholars whose areas of expertise could unambiguously be categorized, a majority were found to have Kripkean intuitions. However, semanticists and philosophers of language (who were likely to be familiar with

⁶ Furthermore, Lam’s attempt to move the debate forward with a set of new experiments is problematic in a number of ways. For example, in his first experiment Lam asked participants the following two questions about a Gödel-style case involving a man named ‘Spencer’ who authors *Romeo and Juliet* and a man named ‘Shakespeare’ who steals credit for it:

Question 1: When these people [who associated the description ‘author of *Romeo and Juliet*’ with the name ‘Shakespeare’] use the name ‘Shakespeare’ in a conversation, is their use of the name to talk about: A. Shakespeare. B. Spencer.

Question 2: When these people use the sentence ‘Shakespeare was English,’ is what they say: A. True. B. Not true.

Lam (2010, 322) suggests, “For both questions, the ‘A’ answers are consonant with the causal-historical theory of names, the ‘B’ answers are consonant with the descriptivist theory of names.” Lam, however, is mistaken in thinking that descriptivists cannot disquote. Regardless of one’s views about reference, one can always say “‘Shakespeare’ refers to Shakespeare.” If descriptivism were true, the referent of the second ‘Shakespeare’ in the previous sentence, and hence of ‘Shakespeare’ in answer A, would simply be determined according to descriptivist principles. Lam makes the same mistake in his second experiment as well.

In Lam’s third experiment, Question 1 becomes the following: ‘When these villagers use the name ‘Shakespeare’ is their use of the name to refer to: A. The real author of *Romeo and Juliet*, Spencer. B. The false author of *Romeo and Juliet*, Shakespeare.’ Lam (2010, 324) claims, “In Question 1, answer A reveals the intuition that the name refers to the actual satisfier of a definite description, and answer B reveals the intuition that the name refers to the person who in fact historically possesses that name.” Lam, however, is ignoring the fact that his question confusingly combines elements that pull in opposite directions. If one wanted simply to disquote, one might focus upon the fact that the name used in ‘B’ is identical to the mentioned name in the question. Alternatively, if one focused upon who is the ‘real’ or the ‘false’ author, one might be led to answer differently. But how likely participants are to choose which of these conflicting factors to follow would tell us nothing about their true semantic intuitions. Because of these problems, Lam’s conclusion that native Cantonese speakers failed to exhibit more descriptivist intuitions than native English speakers cannot be substantiated.

An additional difficulty with Lam’s experiments is that his second questions retain the same ambiguity between speaker’s reference and semantic reference that many authors have noted in Machery et al.’s (2004) original study. Cf. Machery et al. (2010) for further critical discussion of Lam’s experiments.

Kripke's *Naming and Necessity*) were found to be significantly more likely to report Kripkean intuitions (86.4%) than sociolinguists, historical linguists, and anthropological linguists, or discourse analysts (68.7%), whose work inclines them to be sensitive to the descriptions associated with words. These findings strike a critical blow against the Expertise Defense, since expertise was not found to make experts' intuitions consistently more Kripkean.

One of the most significant objections that has been raised to the work of Machery et al. (2004) is the charge that their original probes were ambiguous in ways that threaten to undermine their most important conclusions. As we saw above, the question that originally appeared at the end of Gödel1 is the following:

When John uses the name 'Gödel,' is he talking about:

(A) the person who really discovered the incompleteness of arithmetic? or

(B) the person who got hold of the manuscript and claimed credit for the work?

Ludwig (2007, 150) was one of the first to argue that this question is ambiguous:

For anyone at all familiar with work in the philosophy of language, it is immediately evident that the question does not clearly distinguish between two things: whom John intends to be talking about (or speaker's reference) and who the name John uses refers to, taken literally in the language he intends to be speaking (semantic reference).

Deutsch (2009, 455) has raised the same objection:

In any case, the fact that the vignette question can be interpreted as either (Q1) [viz., "To whom does John intend to refer when he uses 'Gödel'?"], which asks for the speaker's reference of John's uses of 'Gödel', or (Q2) [viz., "To whom does the name 'Gödel' refer when John uses it?"], which asks for the semantic reference of those uses, casts severe doubt on Mallon et al.'s claim that the polls' results show that there are cross-cultural differences in referential

intuitions. Given the ambiguity of the vignette question, it may be that some of their respondents were answering (Q1), while some were answering (Q2).

Similarly, Ichikawa, Maitra, and Weatherson (2012, 59) write, “[I]t’s hard for us to hear a question about who someone is talking about as anything but a question about speaker reference.” Machery and his colleagues agree that the original experiments are theoretically significant only if they reflect participants’ intuitions about semantic reference and not about speaker’s reference because the debate between descriptivism and causal-historical theories concerns the former. Spelling out the upshot of this objection, Machery, Deutsch, and Sytsma (forthcoming) write, “Consequently, it may be that almost everybody has causal-historical intuitions about semantic reference, but that some participants (for some reason, more among Chinese than among Americans) report their intuitions about speaker’s reference.”

Machery, however, is not convinced there is a problematic ambiguity in the original probe:

It makes sense to ask about speaker’s reference only for the occurrences of a proper name, and not for a proper name *qua* type. In addition, to identify the speaker’s reference of the occurrence of a proper name, one needs to know the communicative intention of the speaker....

Now consider the Gödel vignette in Machery et al. (2004). At the end of the vignette, participants are asked who John is talking about “when he uses the name ‘Gödel.’” Since no specific utterance is mentioned and no contextual information is provided that would enable participants to determine John’s communicative intention, it is hard to see how participants could understand the question to be about the speaker’s reference of John’s utterance of ‘Gödel.’ Rather, we submit, on the only plausible interpretation of the question, it is asking about the reference of ‘Gödel’ *qua* type—that is, it is asking about the semantic reference of the term in John’s language. If this is right, then, *pace* Ludwig and Deutsch, the question in the probes is *not* ambiguous. (Machery and Stich 2012, 505, 506)

Furthermore, Machery, Deutsch, and Sytsma (forthcoming) have attempted to settle the matter experimentally by clarifying that the question at the end of the vignette asks about semantic reference rather than speaker's reference. In one experiment Machery, Deutsch, and Sytsma presented Chinese and American participants with Gödel1 and asked them the following, clarified question:

When John uses the name "Gödel," regardless of who he might intend to be talking about, he is actually talking about:

(A) the person who really discovered the incompleteness of arithmetic.

(B) the person who got hold of the manuscript and claimed credit for the work.⁷

This prompt question adds the phrase 'regardless of who he might intend to be talking about' in order to focus participants' attention in the desired direction. Only 38.8% of Chinese participants chose the Kripkean answer (B), compared to 59.5% of American participants who did so. Machery, Deutsch, and Sytsma note that the percentages of East Asian and Western participants choosing the descriptivist answer in response to the clarified probe is similar to the percentages found in a replication that used the original, unclarified probe. In an additional experiment Machery, Deutsch, and Sytsma presented a Chinese translation of Gödel1 and the clarified question above to Chinese participants and found that only 18.5% of them expressed Kripkean intuitions.

In yet another experiment, Machery, Deutsch, and Sytsma (forthcoming) added the following description to Gödel1 in order to make clear that the protagonist has the intention to refer to the

⁷ Machery, Deutsch, and Sytsma (forthcoming) consider the following objection to their clarified probe: "It is sometimes possible to understand "actually talking" as bearing on speaker's reference and "intending to talk about" as bearing on semantic reference, precisely the reverse of the readings we intended in formulating the probe!... We think that this difficulty is not as serious as it might first appear. Although it is true that asking who a person is actually talking about might sometimes lead to a judgment about speaker's reference, there are cases in which it is quite unlikely to do so. In the Clarified Gödel Case in particular, if lay people grasp the distinction between speaker's reference and semantic reference and if our participants paid sufficient attention to the probe question at the end of the Clarified Gödel Case, then it seems unlikely that this probe question could have been understood as being about anything else but semantic reference. One feature of the probe question that supports this assessment is the contrastive nature of the probe question, along with the facts presented to participants in the background story: Our Clarified Gödel Case did not simply ask participants to say who they thought John was actually talking about in using "Gödel," but asked participants to say who they thought John was "actually talking about" *in contrast* to whomever he may be "intending to talk about."

individual who is the semantic reference of ‘Gödel,’ according to causal-historical theories of reference:

One night, John is sitting in his room, reviewing for his mathematics exam by going over the proof of the incompleteness theorem. After a while, he says to his roommate, “Gödel probably got a huge number of awards from mathematical societies!”

In response to this modified vignette, 73.9% of American participants chose the Kripkean answer, but only 56.1% of Chinese participants did so. While this latter percentage is significantly higher than for other samples of Chinese participants, the fact remains that there is a significant difference between the East Asian and Western groups of participants. By showing that cross-cultural differences remain even when questions about reference are clarified, Machery, Deutsch, and Sytsma (forthcoming) provide further evidence that East Asians may well have semantic intuitions that are more in harmony with descriptivist theories of the reference of proper names. Given the lack of unanimity of participant responses within each demographic group, these data also provide further evidence of intra-cultural differences in semantic intuitions.

Thus, it seems that Machery and his collaborators have successfully defused several of the objections that have been raised against their original results, although we think that further exploration and discussion of some of these issues is warranted. In the following sections, we turn our attention to concerns about the work of Machery et al. that have not yet been fully addressed and the studies that we performed to address them.

3. Study 1a: Cross-cultural Differences

One of the things that critics often note about Machery et al.’s (2004) work is that they focus almost exclusively on the results of their Gödel cases and neglect the fact that they failed to find anything of significance with their Jonah cases. Devitt (2011, p. 423), for example, writes:

And whereas their test of Gödel intuitions counts against the refutation [of descriptivism in *Naming and Necessity*], *their test of Jonah intuitions confirmed the refutation*. Yet, it seems to me, the Jonah intuitions are more trustworthy than the Gödel ones because the Jonah case is less fanciful than the Gödel one.

Sytsma and Livengood (2011, 318) concur:

[W]e agree with Ludwig (2007) and Devitt (2011) in thinking that Machery et al. were too quick to embrace the case that indicates variation across culture and to ignore the case that does not.

Although we noted above that there have been successful replications of the initial Gödel results, no replications of the Jonah cases have been reported. Thus, before gathering data using new research materials, we began our exploration of cross-cultural semantic intuitions by trying to replicate all of Machery et al.'s (2004) original results. We also included some measures that we hoped would shed light on intra-cultural variability.

In Study 1, we presented Machery et al.'s (2004) four vignettes to each of 257 undergraduate students of western European ancestry attending state universities in the northeast and midwestern United States (average age = 22, 56% female) and 149 undergraduates attending the University of Hong Kong who were native Chinese speakers (average age = 20.5, 62% female).⁸ The order of the vignettes was counterbalanced. Participants were offered either extra credit in a college course or a chance to win a \$175 gift card. Questionnaires were administered via an online platform hosted by vovici.com.

Due to an error in posting the materials online, the wrong prompt question appeared after the Gödel2 case. Thus, the replication data that we report in this section includes only data from Gödel1, Jonah1, and Jonah2. The data collected from the errant Gödel2 question is described in Section 1c

⁸ An additional 41 American college students of non-western descent and 10 HKU students who were not native Chinese speakers participated in the study but were excluded from the analysis.

below. Participant responses to Gödel1 are summarized in Figure 1, along with Machery et al.'s (2004) original results for the sake of comparison.

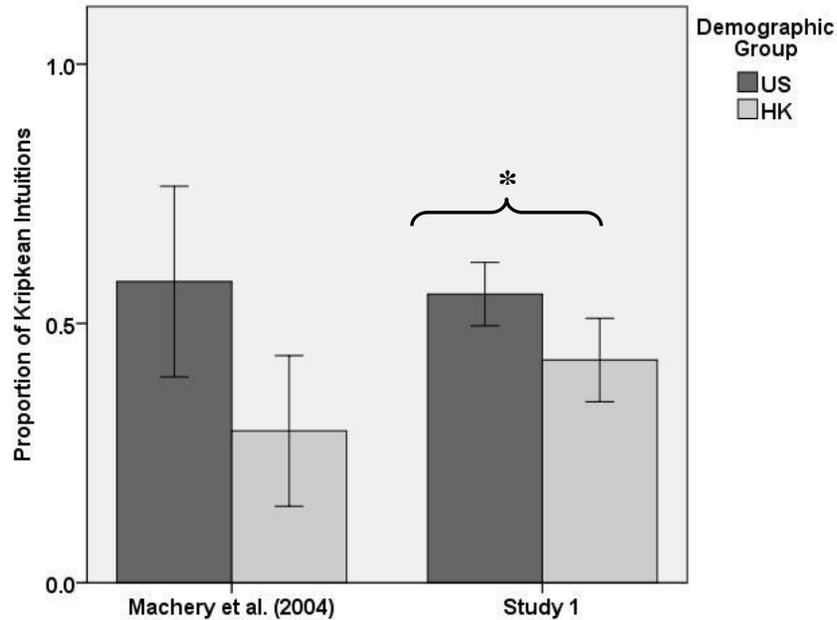


Figure 1. Proportions of Kripkean intuitions given to Gödel1 by Western (.58) and East Asian (.29) participants in Machery et al.'s (2004) original study and by Western (.56) and East Asian (.43) participants in Study 1. Error bars represent 95% confidence intervals in all figures. An ‘*,’ ‘**,’ or ‘***’ indicates a statistically significant difference between pairs of conditions at either the .05, .01, or .001 level.

Like Machery et al. (2004), we observed a significant difference between the proportion of Kripkean intuitions of Western and East Asian participants.⁹ Although the proportion of Kripkean responses given by Western participants in the two studies was almost exactly the same, the proportion of Kripkean responses given by the East Asians in our study was almost 50% higher than in Machery et al.'s. Even though a chi-square test of independence failed to judge this difference to be significant,

⁹ $\chi^2(1, N = 406) = 6.077, p < .05$, Cramér's $V = .12$.

it nonetheless appears large enough to raise some questions about the initial result's ability to reflect the general tendency of Hong Kong students to have descriptivist intuitions about Gödel cases.¹⁰

With the two Jonah cases, 38.9% of Western participants and 25.5% of East Asian participants gave Kripkean responses to Jonah1, while 39.3% of Westerners and 24.8% of East Asians gave Kripkean responses to Jonah2. Because these two pairs of responses were nearly identical, and because Machery et al. (2004) combined responses to both Jonah cases into a single Kripkean intuition score, Figure 2 collapses the percentages we obtained for Jonah1 and Jonah2 and places them beside Machery et al.'s (2004) figures for the sake of comparison.

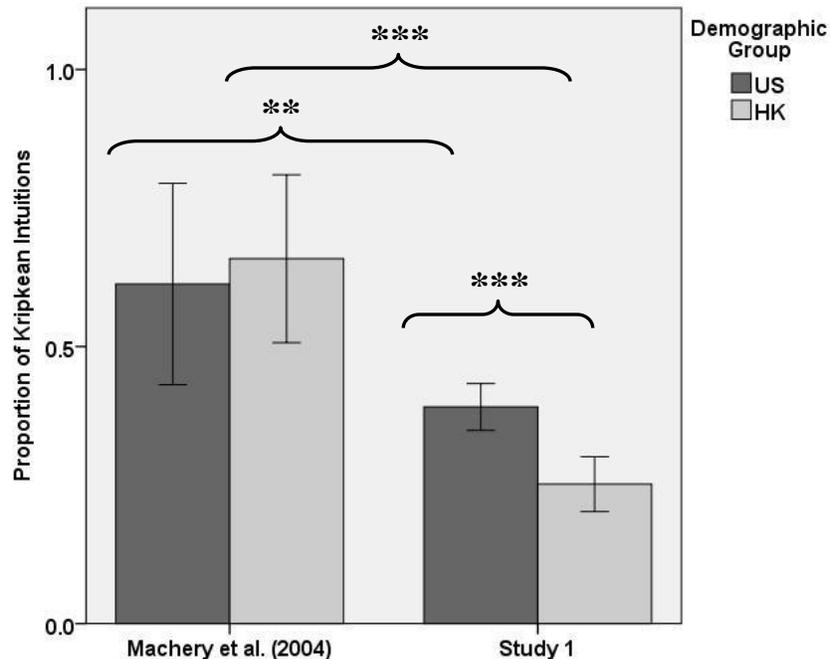


Figure 2. Proportions of Kripkean intuitions given to the Jonah cases by Western (.61) and East Asian (.66) participants in Machery et al.'s (2004) original study and by Western (.39) and East Asian (.25) participants in Study 1.

The data from Study 1 are strikingly different from those of Machery et al. (2004), with both Western and East Asian participants in Study 1 giving significantly fewer Kripkean responses than

¹⁰ $\chi^2(1, N = 190) = 2.509, p > .05$, Cramér's $V = .12$.

Machery et al.'s participants.¹¹ While Machery et al. (2004) found no significant difference between Western and East Asian responses to the Jonah cases, we did find such a difference.¹² Western participants in our study were significantly more likely to give Kripkean responses to the Jonah cases. Notably, both Western and East Asian participants in Study 1 were significantly more likely to give Kripkean responses to Gödel1 than to Jonah1 or Jonah2.¹³

Machery et al. (2004) had predicted that Westerners would be more likely than East Asians to give Kripkean responses to the Jonah cases. Our results may seem to suggest that Machery et al. were right to suspect that cross-cultural differences could be found with the Jonah cases as well. However, a closer look at the data reveals a different story. Machery et al. (2004, B7) offer the following suggestion about why their Jonah results did not turn out the way they expected:

Setting out the Jonah cases precisely requires a lengthy presentation..., so it is possible that our probes were simply too long and complex to generate interpretable data.

Related to the difficulty posed by the rather long and complex structure of the Jonah stories is the fact that Machery et al. did not ask their participants any comprehension questions to see whether they were able to follow the story properly. Therefore, in Study 1 we asked half of our Western and East Asian participants whether the following statements were true or false, in order to test their comprehension of the Jonah cases:

Jonah1a. Attila founded Germany in the second century A.D.

Jonah1b. Raditra founded Germany in the second century A.D.

Jonah2a. Chan Wai Man was a Guangdong nobleman who lived in the mountains around Guangzhou and was forced to live as a thief because of his love for the Minister's daughter.

¹¹ Westerners: $\chi^2(1, N = 576) = 11.217, p < .01$, Cramér's $V = .14$. East Asians: $\chi^2(1, N = 378) = 46.888, p < .001$, Cramér's $V = .35$.

¹² $\chi^2(1, N = 812) = 16.331, p < .001$, Cramér's $V = .14$.

¹³ Westerners: $\chi^2(1, N = 514) = 14.432, p < .001$, Cramér's $V = .17$. East Asians: $\chi^2(1, N = 298) = 10.076, p < .01$, Cramér's $V = .18$.

Jonah2b. Leung Yiu Pang was a Guangdong nobleman who lived in the mountains around Guangzhou and was forced to live as a thief because of his love for the Minister's daughter.

Before reading these comprehension questions, participants were instructed, 'Please answer the following questions on the assumption that the preceding story is true.' According to *Jonah1* and *Jonah2*, none of these statements is true. The other half of our participants did not answer any comprehension questions.

We found that 56% of all participants answered at least one comprehension question incorrectly. Figure 3 depicts the proportions of Kripkean intuitions when participants are separated into those who passed, failed, or did not complete the comprehension task. Note that the 'Did Not Complete' condition includes both participants who would have passed the comprehension test and others who would not have passed.

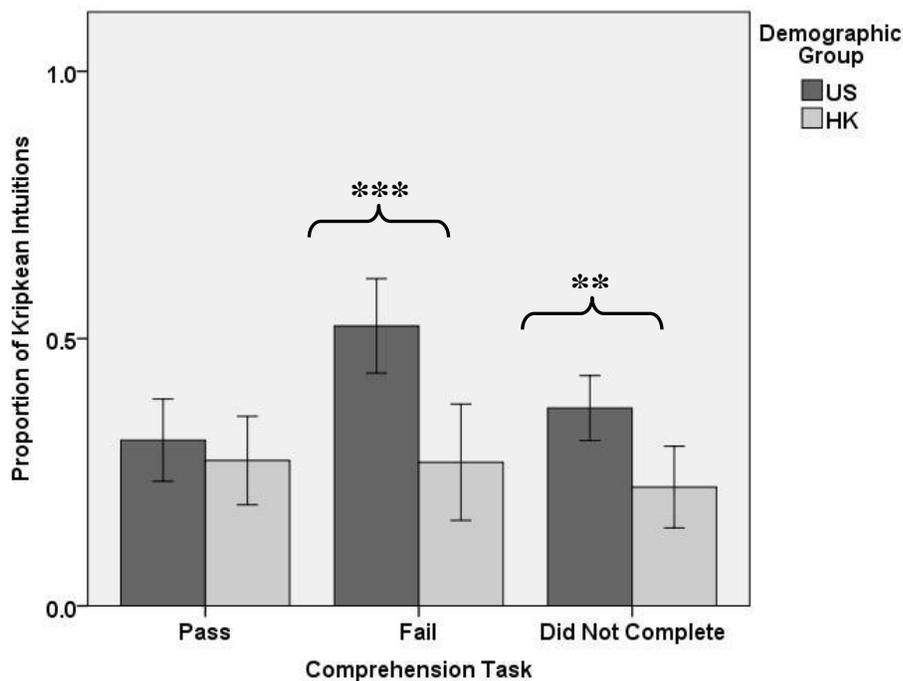


Figure 3. Proportions of Kripkean intuitions given to the *Jonah* cases by Western (.31) and East Asian (.27) participants who correctly answered the comprehension questions correctly, Western (.52) and East Asian (.27) participants who did not answer the comprehension questions correctly, and Western (.37) and East Asian (.22) participants who did not complete the comprehension task in Study 1.

The East Asian responses in the three conditions are nearly equivalent, but the Western responses differ significantly from one another.¹⁴ Thus, what is driving the statistically significant difference we initially observed between Western and East Asian responses to the Jonah cases are the responses of Western participants who did not exhibit a correct understanding of the vignettes. As we will in Section 7 below, we were able to find stable cross-cultural differences in responses to the Jonah cases using different probe questions, even when looking only at the responses of participants who accurately completed a comprehension task. However, in the present study, which used Machery et al.'s original probe questions, we were not able to obtain reliable cross-cultural differences.

It is noteworthy that neither Western nor East Asian responses to Gödel1 or the Jonah cases are as strongly Kripkean as one might expect, if the Kripkean perspective is supposed to provide a correct account of the reference of proper names as they are used in ordinary linguistic practice. Additionally, the high failure rate on our comprehension task suggests that simpler experimental materials should be employed to study folk semantic intuitions.

4. Study 1b: Individual Differences

Although most of the debate about semantic intuitions has focused on the cross-cultural variation found in Machery et al.'s (2004) original study, Machery and his collaborators have always maintained that intra-cultural variation raises important philosophical issues in its own right:

While our focus has been on cultural differences, the data also reveal considerable intra-cultural variation. The high standard deviations in our experiment indicate that there is a great deal of variation in the semantic intuitions within both the Chinese and Western groups. This might reflect smaller intra-cultural groups that differ in their semantic intuitions. A more extreme but

¹⁴ $\chi^2(2, N = 514) = 13.718, p < .01$, Cramér's $V = .16$.

very live possibility is that the variability exists even at the individual level, so that a given individual might have causal-historical intuitions on some occasions and descriptivist intuitions on other occasions. If so, then the assumption of universality is just spectacularly misguided. (Machery et al. 2004, B8)

However, because Machery et al. provide limited information about the contours of the intra-cultural variation within their study, it is difficult to draw any firm conclusions about it from their study. In order to shed light on this issue in our own study, we computed a Kripkean intuition score for each of our participants and compared the variability of this measure across individuals. Assigning a ‘1’ for each Kripkean response to Gödel1, Jonah1, and Jonah2 results in there being four possible Kripkean intuition scores (from 0 to 3). In both the Western and East Asian participant populations, the median and mode scores were 1 (cf. Figure 4).

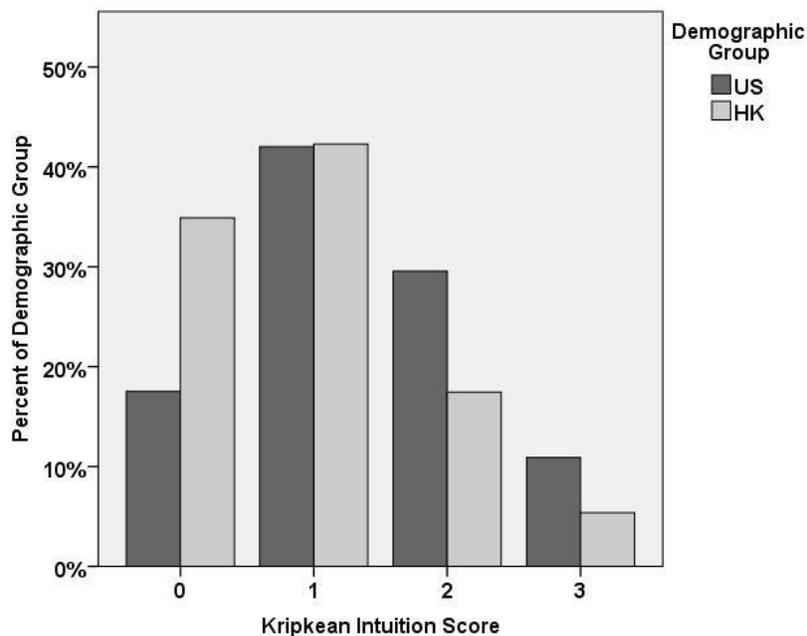


Figure 4. Percentages of Western and East Asian participants with different Kripkean intuition scores in Study 1.

42% of Western and 42% of East Asian participants had a Kripkean intuition score of 1, with 60% of Westerners and 75% of East Asians having a score of either 0 or 1. The mean Western score was 1.34, while the mean East Asian score was .93. This difference is statistically significant.¹⁵ When the analysis of Kripkean intuition scores is restricted to only those participants who correctly answered all four comprehension questions for the Jonah cases, the overall distribution of the data is relatively unchanged—the mean, median, and mode responses remain centered around 1. Thus, we see that neither participant group was terribly Kripkean in their responses to the test materials. This fact, together with the degree of spread of participant responses across all four possible scores seems to provide support for Machery et al.’s challenge to Kripkean orthodoxy and to any claims about the alleged uniformity of semantic intuitions.

In order to investigate intra-cultural variation in greater detail, we also directed half of the participants in Study 1 to complete a ten-item personality inventory developed by Samuel Gosling, Peter Rentfrow, and William B. Swann Jr. (2003). The inventory was a brief measure of the Big Five personality traits of extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience, which together provide the most widely accepted framework for assessing the broadest features of human personality.¹⁶ We wanted to see whether the degree of intra-cultural variation in semantic intuitions between those whose personalities differ along these lines might rival the degree of cross-cultural variation observed in Machery et al. (2004) and in Study 1.

We found significant positive correlations between participants’ Kripkean intuition scores and their self-reported conscientiousness ($r = .20, p < .01$), emotional stability ($r = .15, p < .05$), and openness to experience ($r = .16, p < .05$). Conscientiousness (sometimes called ‘will to achieve’) is a cluster of traits that includes self-discipline, efficiency, organization, and dependability. Emotional

¹⁵ $t(404) = 4.476, p < .001, r = .22$.

¹⁶ As Nettle (2007, 9) describes it, “The five-factor model has emerged from a welter of research over the last few decades and looks to be the most comprehensive, reliable and useful framework for discussing human personality that we have ever had.”

stability concerns how easily individuals experience negative emotions, such as anger, anxiety, depression, and vulnerability. Openness to experience is a combination of traits such as curiosity, inventiveness, creativity, preference for novelty, and emotional self-awareness. Each of these three traits has been shown to be positively correlated with academic achievement (Komarraju et al. 2009; Komarraju et al. 2011). Thus, defenders of Kripkean orthodoxy might find a measure of encouragement in the fact that individuals who are more likely to be better students were also more likely to report Kripkean intuitions. However, the most important point for our purposes is that the correlations between these personality traits and participants' Kripkean intuition scores were almost as strong as the association between being Western and having a higher Kripkean intuition score ($r = .22$, $p < .001$).¹⁷ We conclude that if the cross-cultural differences reported by Machery et al. (2004) pose an interesting philosophical challenge, the individual differences we have observed should do so as well.

5. Study 1c: Clarified Perspectives

Justin Sytsma and Jonathan Livengood (2011, 319-320) have recently raised a potentially serious objection to the work of Machery et al. (2004). They argue that, in addition to there being an ambiguity between speaker's reference and semantic reference in the original probes for the Gödel cases, there is a second and more fundamental ambiguity in the perspective from which participants should understand the materials:

First, the question used in Machery et al.'s Gödel probe does not clearly indicate whether the (A) and (B) answer choices are to be read from the *narrator's epistemic perspective* (the narrator relaying information of which John is ignorant) or rather from *John's epistemic*

¹⁷ Despite the fact that there are several studies that support the cross-cultural applicability of the Big Five personality model (e.g., McCrae et al., 1996; McCrae & Costa, 1997; Rolland, 2002; McCrae et al., 2005), in our study there were significant correlations between being Western and reporting higher conscientiousness ($r = .4$, $p < .001$), emotional stability ($r = .25$, $p < .001$), and openness to experience ($r = .41$, $p < .001$).

perspective (as the speaker using the name ‘Gödel’)... While Machery et al. expect the descriptions to be read from the narrator’s perspective, the question might plausibly lead participants to instead adopt John’s perspective.... Specifically, from the narrator’s point of view, ‘the person who really discovered the incompleteness of arithmetic’ denotes Schmidt and ‘the person who got hold of the manuscript and claimed credit for the work’ denotes Gödel; but, from John’s perspective ‘the person who really discovered the incompleteness of arithmetic’ denotes Gödel. Indeed, John has never heard of Schmidt!

As with the ambiguity between semantic reference and speaker’s reference, if participant responses differ because members of one demographic group are more likely to choose one disambiguation than participants in another group, the resulting data will tell us nothing about their true semantic intuitions.

In order to emphasize the narrator’s perspective, Machery et al. (2004) used the adverb ‘really’ in the first answer to the Gödel1 prompt (‘the person who really discovered the incompleteness of arithmetic’), but Sytsma and Livengood (2011, n. 6) find this to represent “insufficient emphasis.” Consequently, they constructed two additional probes that clarified whether participants should adopt John’s perspective or the narrator’s:

John’s Perspective: When John uses the name ‘Gödel,’ does John think he is talking about: (A) the person who the story says really discovered the incompleteness of arithmetic? Or, (B) the person who the story says got hold of the manuscript and claimed credit for the work?

Narrator’s Perspective: When John uses the name ‘Gödel,’ is he actually talking about: (A) the person who the story says really discovered the incompleteness of arithmetic? Or, (B) the person who the story says got hold of the manuscript and claimed credit for the work?

Sytsma and Livengood presented Western subjects with Machery et al.’s original probe and the two clarified probes in a between-subjects design and obtained the results represented on the left-hand side of Table 4.

	Sytsma & Livengood's Westerners	Machery et al.'s Westerners	Machery et al.'s East Asians
Original probe	39.4%	58%	29%
John's perspective	22%	n/a	n/a
Narrator's perspective	57.4%	n/a	n/a

Table 4. Percentages of Kripkean intuitions given by Western and East Asian participants to Gödel in Sytsma & Livengood's (2011) and Machery et al.'s (2004) studies.

Significantly more participants in Sytsma and Livengood's study reported Kripkean intuitions when the narrator's perspective was highlighted than when John's perspective was highlighted. The percentage of Kripkean intuitions given in response to the original probe is almost exactly equal to the arithmetic mean of the other two categories of responses. Note, too, that the 39.4% who chose the Kripkean answer in response to the original probe represent a failure to replicate Machery et al.'s original finding that Westerners were rather likely to express Kripkean intuitions in response to Gödel cases. Indeed, there was no statistically significant difference between Sytsma and Livengood's westerners and Machery et al.'s (2004) East Asians.

In an additional experiment, Sytsma and Livengood presented participants with a probe that clarified to an even further extent that the narrator's perspective should be adopted when thinking about Gödel:

Clarified Narrator's Perspective: Having read the above story and accepting that it is true, when John uses the name 'Gödel,' would you take him to actually be talking about: (A) the person who (unbeknownst to John) really discovered the incompleteness of arithmetic? Or, (B)

the person who is widely believed to have discovered the incompleteness of arithmetic, but actually got hold of the manuscript and claimed credit for the work?

A large percentage of ordinary participants (73.8%) chose the Kripkean answer (B). Because significantly different percentages of participants chose Kripkean answers when different epistemic perspectives were highlighted, Sytsma and Livengood (2011, 328) conclude:

While our data are specific to Westerners, they show that Machery et al.'s instrument is broken... [W]e argue that our results (1) undermine Machery et al.'s claim to have shown significant variation in semantic intuitions *within* Westerners, (2) undermine their claim to have shown significant variation in semantic intuitions *cross-culturally* between Westerners and East Asians, and (3) undermine their claim to have shown significant variation in semantic intuitions *within* East Asians.

In spite of these interesting results, we suspected that it was nonetheless likely that Western and East Asian participants would respond differently to the clarified narrator's perspective probe. When we presented the participants in Study 1 with the clarified narrator's perspective probe in conjunction with Gödel², that is precisely what we found. Significantly fewer East Asians gave Kripkean responses than Westerners (cf. Figure 5).¹⁸

¹⁸ $\chi^2(1, N = 406) = 8.362, p < .01$, Cramér's V = .14.

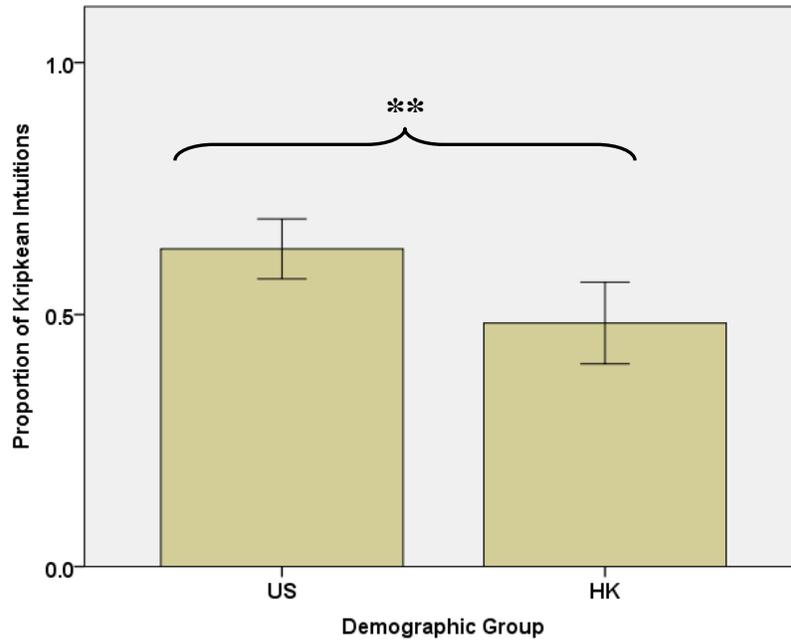


Figure 5. Proportions of Kripkean responses to the clarified narrator’s perspective version of Gödel2 among Western (.63) and East Asian (.48) participants.

Although we used Machery et al.’s original, unclarified prompt question for Gödel1 and Sytsma and Livengood’s clarified prompt question for Gödel2, neither our Western nor our East Asian participants responded in a significantly different manner to the two cases. We thus have a failure to replicate Sytsma and Livengood’s results.

Sytsma and Livengood (2011, 328) contend that their results cast doubt on intra-cultural variation among Western participants because of the high rate of Kripkean responses they obtained with their ‘clarified narrator’s perspective’ probe:

Our findings are most directly at odds with the claim that there is significant variation in semantic intuitions within Westerners. Western participants find Machery et al.’s original test question to be ambiguous. Specifically, alterations to the question that should be irrelevant to participants’ semantic intuitions significantly alter their responses: clarifying the question to emphasize the narrator’s perspective produced a dramatic increase in (B) [i.e., Kripkean] answers.

However, the percentage of Western participants that we observed giving Kripkean responses to Gödel2 was only 63%—a figure that is low enough to raise questions about intra-cultural uniformity. Furthermore, the East Asian responses were split roughly in half, suggesting significant intra-cultural variation within this group. While Sytsma and Livengood may well have pinpointed an ambiguity in Machery et al.’s original probe, the fact remains that both cross-cultural and intra-cultural differences in semantic intuitions can be found using clarified probes.

6. Study 2a: Shifting Anchors

An important feature of Gödel1 is that it tells participants one thing about the bearer of the name ‘Gödel’ but then tells them to suppose that something else is true. The vignette begins, “Suppose that John has learned in college that Gödel is the man who proved an important mathematical theorem, called the incompleteness of arithmetic.” ‘Learns’ is commonly taken by philosophers to be factive, such that learning that p entails p . Thus, the initial sentence of Gödel1 seems to presuppose Gödel really is the discoverer of the theorem. However, a few sentences later participants are told, “Now suppose that Gödel was not the author of this theorem.”¹⁹ We hypothesized that the initial description might have an anchoring effect and that participants who would otherwise be inclined to offer up descriptivist intuitions might ‘adjust’ away from the initial anchor less than they should (as is common in such cases), given their true semantic intuitions (cf. Tversky & Kahneman, 1974).

Therefore, we constructed the following modified versions of Gödel1 and Gödel2 that began by associating the name ‘Schmidt’ with the man who proved the incompleteness theorem and the name ‘Zhang Shoujing’ with the astronomer who first determined the summer and winter solstices, leaving other details of the stories the same.

¹⁹ Lam (2010, 322) does the same thing, having one vignette begin “Suppose there is a group of people who do not know anything of the English author Shakespeare except the name and that he is the author of ‘Romeo and Juliet.’” If they know this one thing about Shakespeare, then it must be true. But participants are then told “Unbeknownst to this group of people, Shakespeare did not in fact write the play ‘Romeo and Juliet.’”

*Gödel*³. Suppose that there was a man called ‘Schmidt’ who proved an important mathematical theorem called the incompleteness of arithmetic and that Schmidt’s body was found in Vienna under mysterious circumstances many years ago. Suppose, however, that Schmidt’s friend Gödel somehow got hold of Schmidt’s manuscript and claimed credit for the work, which was thereafter attributed to Gödel. Now suppose that John has learned in college that Gödel is the man who proved the incompleteness of arithmetic. John is quite good at mathematics and he can give an accurate statement of the incompleteness theorem, which he attributes to Gödel as the discoverer. But this is the only thing that he has heard about Gödel. Most people who have heard the name ‘Gödel’ are like John; the claim that Gödel discovered the incompleteness theorem is the only thing they have ever heard about Gödel.

*Gödel*⁴. Suppose that the precise time of the summer and winter solstices was first determined by an astronomer named ‘Zhang Shoujing’ who died soon after making the discovery. Suppose, however, that credit for the discovery was stolen by someone named Tsu Ch’ung Chih. Suppose that the theft remained entirely undetected and that Tsu Ch’ung Chih became famous for the discovery of the precise times of the solstices. Ivy is a high-school student in Hong Kong. In her astronomy class she was taught that Tsu Ch’ung Chih was the man who first determined the precise time of the summer and winter solstices. Like all her classmates, this is the only thing she has heard about Tsu Ch’ung Chih. Many people are like Ivy; the claim that Tsu Ch’ung Chih determined the solstice times is the only thing they have heard about him.

We retained the original questions that appeared after Gödel¹ and Gödel².

161 undergraduate students of western European ancestry attending state universities in the northeast and midwestern United States (average age = 20, 53% female) and 46 undergraduates attending the University of Hong Kong whose native language was Chinese (average age = 20, 70% female) participated in Study 2 in exchange for either extra credit in a college course or a chance to

win a \$175 gift card. Questionnaires were administered via an online platform hosted by vovici.com.²⁰

The order of the vignettes was counterbalanced.

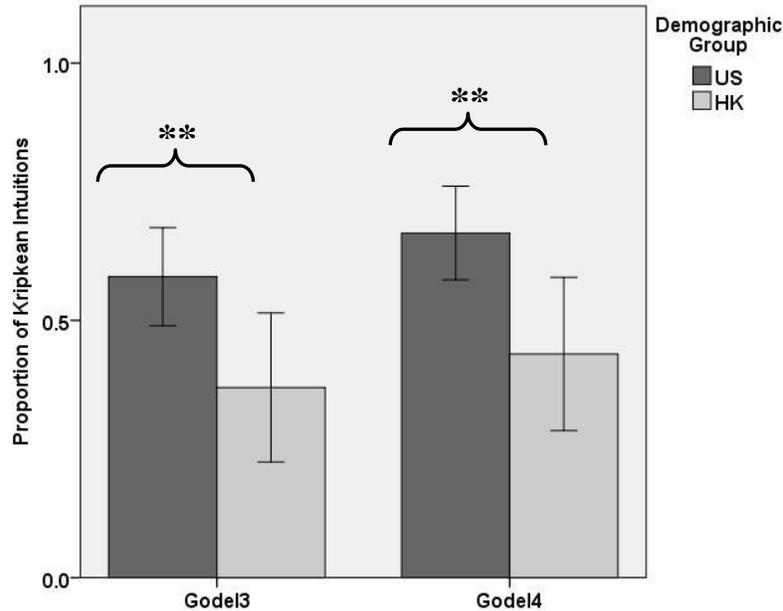


Figure 6. Proportions of Kripkean responses to Gödel3 (US = .58, HK = .37) and Gödel4 (US = .67, HK = .43) in Study 2a.

Again, we found cross-cultural differences in semantic intuitions. In both Gödel3 and Gödel4, Westerners were more likely than East Asians to give Kripkean responses.²¹ East Asian responses to Gödel1, Gödel3, and Gödel4 did not differ significantly from one another. Western responses to Gödel3 did not differ significantly from Western responses to Gödel1 or Gödel4, but Western responses to Gödel4 were slightly higher than those for Gödel1.²² Although this difference is statistically significant, we do not think it is theoretically significant because of the consistency of the other results. Thus, one more potential worry for the original findings of Machery et al. proves to be toothless, and cross-cultural differences in semantic intuitions prove to be robust.

²⁰ An additional 20 American college students of non-western descent and 4 HKU students who were not native Chinese speakers participated in the study but were excluded from the analysis.

²¹ Gödel 3: $\chi^2(1, N = 207) = 7.845, p < .01$, Cramér's $V = .20$. Gödel4: $\chi^2(1, N = 152) = 7.375, p < .01$, Cramér's $V = .22$.

²² Gödel1 vs. Gödel4: $\chi^2(1, N = 363) = 3.988, p < .05$, Cramér's $V = .11$.

7. Study 2b: Modified Jonah Probes

In addition to speculating that the long and complex structure of their Jonah cases may have been partly responsible for their failure to find any cross-cultural differences in intuitive responses to them, Machery et al. (2004, B7) offered the following suggestion:

[I]n the Jonah cases, the descriptivist response is that the speaker's term fails to refer. It might be that for pragmatic reasons, both the Westerners and the Chinese reject the uncharitable interpretation that the speaker is not talking about anyone.

Recall that the question Machery et al. asked participants about Jonah1 was the following:

When a contemporary German high-school student says "Attila was the king who drove the Romans from Germany," is he actually talking about the wise and gentle nobleman, Raditra, who is the original source of the Attila legend, or is he talking about a fictional person, someone who does not really exist?

(A) He is talking about Raditra.

(B) He is talking about a fictional person who does not really exist.

Machery et al. conjectured that charity might lead participants to try to find a person to whom the speaker might be referring. In addition, the very idea of talking about things that aren't things at all (because they do not exist) might cause some processing difficulties with philosophically untrained participants.

Henry Jackman (2009, 169-170) expresses related concerns about these probes:

Further, it should also be noted that when [Machery et al.] present the Jonah case to their informants, what they treat as the 'descriptivist' doesn't posit reference failure as much as it does the speaker referring to a fictional character. When [Machery et al.] outline the descriptive theory it involves a clear commitment to the effect that "If the description is not satisfied at all

or if many individuals satisfy it, the name does not refer,” but when they present the case, this commitment disappears, and the name is taken to refer to “a fictional person who does not really exist.

Consequently, in Study 2 we constructed probes that we thought might both prevent considerations of charity from having a significant effect and allow genuine reference failure to be an option.

After reading each of the original Jonah cases and answering the comprehension questions described above, participants were directed to answer the following questions about their semantic intuitions:

Jonah3. When a contemporary German high-school student says “Attila was the king who drove the Romans from Germany,” is he actually talking about the wise and gentle nobleman, Raditra?

(A) He is talking about Raditra.

(B) He is not talking about Raditra.

Jonah4. When Mei Ling says “Chan Wai Man stole from the rich and gave to the poor,” is she actually talking about the generous monk, Leung Yiu Pang?

(A) She is talking about the generous monk, Leung Yiu Pang.

(B) She is not talking about the generous monk, Leung Yiu Pang.

Results combining participant responses to Jonah3 and Jonah4 are summarized in Figure 7.

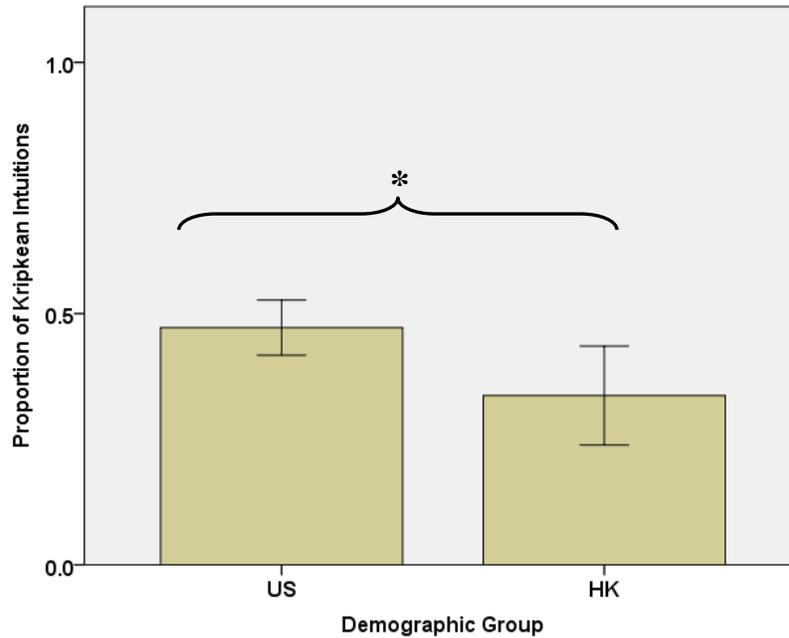


Figure 7. Proportions of Kripkean intuitions given to the Jonah cases by Western (.47) and East Asian (.34) participants in Study 2.

Significantly more Westerners than East Asians gave Kripkean answers to the key questions about the reference of the name ‘Jonah.’²³ Most importantly, when the results are restricted to those participants who answered the comprehension questions correctly, a significant difference remains.²⁴ Westerners who correctly completed the comprehension tasks gave Kripkean responses 42% of the time, and 28% of the relevant subset of East Asians did so as well. Thus, contrary to the results of Machery et al. (2004) and Study 1, significant cross-cultural differences can be found in Jonah cases after all. However, because these results have not always been found, they may be less robust than those associated with Gödel cases.

²³ $\chi^2(1, N = 414) = 5.295, p < .05$, Cramér’s $V = .11$.

²⁴ $\chi^2(1, N = 254) = 4.241, p < .05$, Cramér’s $V = .13$.

8. Study 2c: Intra-Cultural Variation

Finally, we computed Kripkean intuition scores for each of the participants in Study 2. The distribution of these scores is represented in Figure 8.

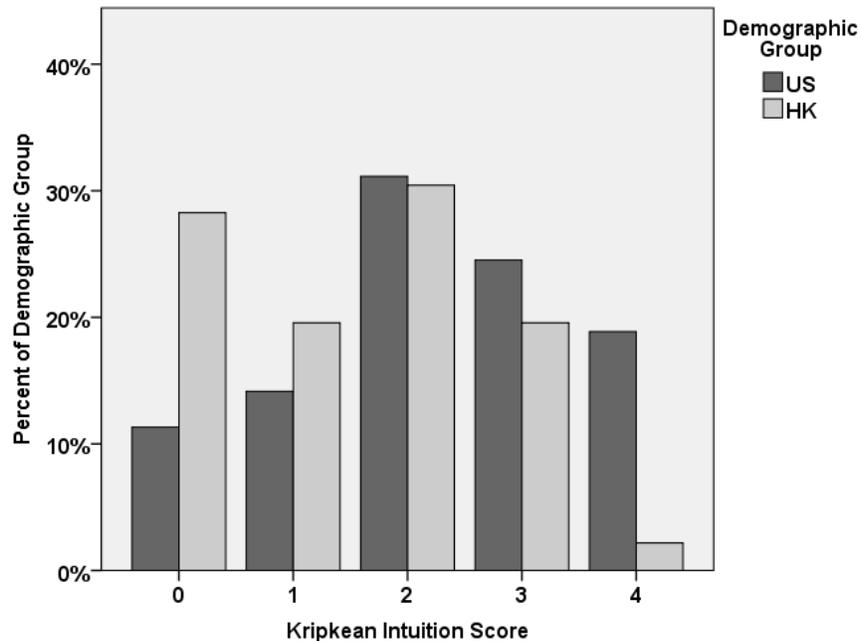


Figure 8. Percentages of Western and East Asian participants with different Kripkean intuition scores in Study 2.

For both demographic groups, the mode and median responses are 2. The mean Kripkean intuition score for Westerners is 2.25, while the mean for East Asians is 1.48. This difference is significant.²⁵ In both Study 1 and Study 2, the distribution of Kripkean intuition scores approximates the standard normal distribution to a reasonable degree, whereas the truth of causal-historical theories of reference might lead one to expect this distribution to be skewed more toward higher scores.

²⁵ $t(150) = 3.602, p < 001, r = .28.$

9. Discussion

In our experience, when the results of Machery et al. (2004) are mentioned in philosophical conversation, the most common response is for those present to comment on how deeply flawed they think the study is and how unreliable they find the results to be. However, subsequent empirical research (including the studies reported above) seems to vindicate the original findings. Cross-cultural differences in responses to Gödel cases remain when additional variables are controlled, and cross-cultural differences in responses to Jonah cases can sometimes be found. Furthermore, when we turn our attention to individual differences, we see that the range of variability within cultures typically outpaces the variability between them.

We think these findings pose a challenge to work in the philosophy of language that is premised on the uniformity conjecture, which Sytsma and Livengood (2011, p. 317) describe as follows:

The uniformity conjecture then has two parts. The first part claims that there exists widespread agreement among the intuitive judgments people make with respect to philosophical thought experiments or case studies concerning reference. Universal agreement would be an unreasonably strong requirement. With that exception, we will leave vague just how widespread the agreement needs to be.... The second part of the uniformity conjecture claims that agreement persists if one conditions on membership in theoretically interesting groups.

We think the challenge to this conjecture is analogous to the challenge that would be posed to the epistemological tradition if the folk were found to regularly attribute knowledge to subjects whose beliefs were not true. Since epistemologists generally take themselves to be explicating the ordinary conception of knowledge and have always maintained that knowledge entails truth, an interesting philosophical challenge would arise if the habits of those who competently use the concept of knowledge depart from how philosophers say they both do and should apply the concept.

One common response to philosophical challenges raised by studies of folk intuitions is to dismiss the intuitive responses of the unsophisticated and untrained as unreliable. As we noted above, however, Machery (2012) found that even among linguists and philosophers of language there were differences in how Kripkean their responses were. A different kind of response to our studies and those of Machery and his collaborators is to try to find some way to accommodate the observed variability. For example, when James Genone and Tonia Lombrozo (2012) found that individuals were not consistent in relying upon descriptive or causal information when making judgments about the reference of natural kind terms, they argued that neither a pure descriptive theory nor a pure causal-historical theory would suffice and proposed a hybrid theory of reference instead. Moreover, they argued that the variation they observed could not be accommodated by appeal to a mixed population of pure descriptive theorists and pure causal theorists.

We do not have the space to adjudicate between competing explanations of the individual and cross-cultural differences in semantic intuitions we have found. Our primary goal has been to report the results of new empirical studies that control for variables that some critics have expressed concern about, to situate those results within the larger debate about cross-cultural semantic intuitions, and to demonstrate the robustness of the individual and cross-cultural differences that have been found. We conclude that critics can no longer dismiss Machery et al.'s (2004) original findings as flukes or as the result of poor experimental controls.

References

- Deutsch, M. 2009. Experimental philosophy and the theory of reference. *Mind & Language*, 24, 445-66.
- Devitt, M. 2011. Experimental semantics. *Philosophy and Phenomenological Research* 82: 418-35.
- Genone, J., and T. Lombrozo. 2012. Concept possession, experimental semantics, and hybrid theories of reference. *Philosophical Psychology*, 25: 717-42.

- Gosling, S. D., Rentfrow, P. J., and Swann, W. B., Jr. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37, 504-28.
- Ichikawa, J., I. Maitra, and B. Weatherson. (2012). In defense of a Kripkean dogma. *Philosophy and Phenomenological Research*, 85, 56-68.
- Jackman, H. 2009. Semantic intuitions, conceptual analysis, and cross-cultural variation. *Philosophical Studies*, 146, 159-77.
- Kallestrup, J. in progress. Counterfactuals, counterfactuals and semantic intuitions.
- Komarraju, M., Karau, S. J., and Schmeck, R. R. 2009. Role of the Big Five personality traits in predicting college students' academic motivation and achievement. *Learning and Individual Differences*, 19, 47-52.
- Komarraju, M., Karau, S. J., Schmeck, R. R., and Avdic, A. 2011. The Big Five personality traits, learning styles, and academic achievement. *Personality and Individual Differences*, 51, 472-7.
- Kripke, S. 1972/1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Lam, B. 2010. Are Cantonese-speakers really descriptivists? Revisiting cross-cultural semantics. *Cognition* 115, 320-29.
- Ludwig, K. 2007. The Epistemology of Thought Experiments: First Person versus Third Person Approaches. *Midwest Studies in Philosophy* 31, 128-59.
- Machery, E. 2011. Variation in intuitions about reference and ontological disagreements. In S. D. Hales (ed.), *A Companion to Relativism*. New York: Blackwell, pp. 118-36.
- Machery, E. 2012. Expertise and intuitions about reference. *Theoria* 27: 37-54.
- Machery, E., M. Deutsch, R. Mallon, S. Nichols, J. Sytsma, and S. Stich. 2010. Semantic intuitions: Reply to Lam. *Cognition* 117: 361-6.
- Machery, E., M. Deutsch, J. Sytsma. forthcoming. Speaker's reference and cross-cultural semantics.
- Machery, E., R. Mallon, S. Nichols, and S. Stich. 2004. Semantics, cross-cultural style. *Cognition*, 92, B1-B12.
- Machery, E., R. Mallon, S. Nichols, and S. Stich. 2013. If folk intuitions vary, then what? *Philosophy and Phenomenological Research*, 86, 618-83.

- Machery, E., C. Y. Olivola, and M. de Blanc. 2009. Linguistic and metalinguistic intuitions in the philosophy of language. *Analysis*, 69, 689-94.
- Machery, E., and S. Stich. 2012. Experimental philosophy of language. In G. Russell & D. G. Fara (Eds.), *Routledge Companion to the Philosophy of Language*. New York: Routledge, pp. 495-512.
- Mallon, R., E. Machery, S. Nichols, and S. Stich. 2009. Against arguments from reference. *Philosophy and Phenomenological Research*, 79: 332-56.
- Martí, G. 2009. Against semantic multi-culturalism. *Analysis* 69: 42-8.
- Martí, G. forthcoming. Empirical data and the theory of reference. In M. O'Rourke (ed.), *Topics in Contemporary Philosophy*, vol 10: Reference and Referring. Cambridge, MA: MIT Press.
- McCrae, R. R., Costa, P. T., Jr., and Yik, M. S. M. 1996. Universal aspects of Chinese personality structure. In M. H. Bond (ed.), *The Handbook of Chinese Psychology*. Hong Kong: Oxford University Press, pp. 189-207.
- McCrae, R. R., and Costa, P. T., Jr. 1997. Personality trait structure as a human universal. *American Psychologist*, 52, 509-16.
- McCrae, R. R., Terracciano, A., and 78 Members of the Personality Profiles of Cultures Project. 2005. Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology*, 88, 547-61.
- Nettle, D. 2007. *Personality: What Makes You The Way You Are*. New York: Oxford University Press.
- Rolland, J.-P. 2002. The cross-cultural generalizability of the five-factor model of personality. In R. R. McCrae & J. Allik (eds), *The Five-Factor Model of Personality Across Cultures*. New York: Kluwer Academic/Plenum Publishers, pp. 7-28.
- Sytsma, J., and J. Livengood. 2011. A new perspective concerning experiments on semantic intuitions. *Australasian Journal of Philosophy*. 89: 315-32.
- Tversky, A., and Kahneman, D. 1974. Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-30.

Appendix

*Gödel*². Ivy is a high-school student in Hong Kong. In her astronomy class she was taught that Tsu Ch'ung Chih was the man who first determined the precise time of the summer and winter solstices. But, like all her classmates, this is the only thing she has heard about Tsu Ch'ung Chih. Now suppose that Tsu Ch'ung Chih did not really make this discovery. He stole it from an astronomer who died soon after making the discovery. But the theft remained entirely undetected and Tsu Ch'ung Chih became famous for the discovery of the precise times of the solstices. Many people are like Ivy; the claim that Tsu Ch'ung Chih determined the solstice times is the only thing they have heard about him. When Ivy uses the name 'Tsu Ch'ung Chih,' is she talking about:

(A) the person who really determined the solstice times? or

(B) the person who stole the discovery of the solstice times?

*Jonah*². Lau Mei Ling is a high-school student in the Chinese city of Guangzhou. Like everyone who goes to high-school in Guangzhou, Mei Ling believes that Chan Wai Man was a Guangdong nobleman who had to take refuge in the wild mountains around Guangzhou in the eleventh century A.D., because Chan Wai Man was in love with the daughter of the ruthless Government Minister Lee, and the Minister did not approve. Everyone in Lau Mei Ling's high-school believes that Chan Wai Man had to live as a thief in the mountains around Guangzhou, and that he would often steal from the rich allies of the Minister Lee and distribute their goods to the poor peasants.

Now suppose that none of this is true. No Guangdong nobleman ever lived in the mountains around Guangzhou, stealing from the wealthy people to help the peasants. The real facts are the following. In one of the monasteries around Guangzhou, there was a helpful monk called 'Leung Yiu Pang.' Leung Yiu Pang was always ready to help the peasants around his monastery, providing food in the winter, giving medicine to the sick and helping the children. Because he was so kind, he quickly became the main character of many stories. These stories were passed on from one generation of peasants to the next. Over the years, the story changed slowly as the peasants would forget some elements of the story and add other elements. In one version, Leung Yiu Pang was described as a rebel fighting Minister Lee. Progressively the story came to describe the admirable

deeds of a generous thief. By the late fourteenth century, the story was about a generous nobleman who was forced to live as a thief because of his love for the Minister's daughter. At length, not a single true fact remained in the story.

Meanwhile, the name 'Leung Yiu Pang' was slowly altered: it was successively replaced by 'Cheung Wai Pang' in the twelfth century, 'Chung Wai Man' in the thirteenth, and finally by 'Chan Wai Man.' The story about the adventurous life of Chan Wai Man was written down in the fifteenth century by a scrupulous historian, from whom all our beliefs are derived. Of course, Mei Ling, her classmates and her parents know nothing about these real events. Mei Ling believes a story about a generous thief who was fighting against a mean minister.

When Mei Ling says "Chan Wai Man stole from the rich and gave to the poor," is she actually talking about the generous monk, Leung Yiu Pang, who is the original source of the legend about Chan Wai Man, or is she talking about a fictional person, someone who does not really exist?

- (A) She is talking about the generous monk, Leung Yiu Pang.
- (B) She is talking about a fictional person who does not really exist.