Self-Perception of Understanding and the Division of Linguistic Labor

Abstract

In this paper we set out to determine how self-perception of understanding of a word is affected by the Division of Linguistic Labor (DLL). We developed a between-subjects experiment in which participants were required to draw distinctions between two pairs of highly confusable dog breeds. In the treatment group, participants were grouped into dyads and received complementary training (so that each participant became an expert at identifying the two confusable breeds of only one kind of dog), and were allowed to interact with each other by means of the labels of these four dog breeds. We succeeded in capturing several aspects of DLL found in the broader social world. The results of the experiment show that people's self-perceived understanding was influenced not only by their accuracy at recognizing a word's reference, but also by their perceived knowledge relative to others.

Keywords: community of knowledge; semantic externalism, extended cognition, epistemic dependence, experimental philosophy.

1 Introduction

In his seminal paper *The Meaning of 'Meaning'* (Putnam, 1975), Hilary Putnam presented the thesis that the meanings of some of a speaker's words cannot be in their head, because they might not have the necessary knowledge to recognize the objects referred to by these words. However, these words mean what they mean even when uttered by a "novice" speaker, and Putnam suggests that this is because such a speaker may delegate the recognition task to other speakers in their community who do possess the appropriate knowledge: the "experts." This division between experts and novices is known as the Division of Linguistic Labor (DLL).[1]

[1] Putnam is also associated with the view known as *natural kind externalism*, which entails paradoxical consequences for DLL (cf. Talmage, 1998; Briscoe, 2006). This is not the place to delve into this discussion but suffice it to say that DLL can be dissociated from the notions presupposed by *natural kind externalism* (cf. Ware, 1978; Kallestrup, 2012).

Putnam is also attributed with the view that experts and novices possess the same linguistic skills, since they are competent speakers of the same linguistic community. The difference between them is not that one of them is more linguistically competent than the other, but rather that the expert knows more about the word's reference than does the novice (Putnam, 1975, pp.166–8). Some commentators, such as Briscoe and Talmage (Briscoe, 2006; Talmage, 1998), have shown that Putnam's view is compatible with the idea that the novice possesses a sort of 'world-involving understanding'. In a sense, the novice's thoughts about the objects being referred to by the word depend on the meaning of this word, which in turn are not solely determined by one individual's mental contents, but are also determined by their linguistic community. This view contradicts a more traditional view about thoughts. For instance, Evans (Evans, 1982) propounds Russell's Principle, according to which a person cannot have thoughts about an object that they cannot differentiate from all other objects. In this latter view, the novice cannot have thoughts about the word's reference, since they do not have enough knowledge about it. Clearly, this view contradicts Briscoe's and Talmage's because it entails that the novice cannot possess this so called world-involving understanding (Wagner, 2020).

It is a moot point to try and argue in favor of one or the other view, since they both depend on their idiosyncratic definitions of what a thought is. The issue remains, however, how we can *empirically assess* the extent to which experts and novices understand the meaning of words. The most natural way to address this issue is to directly ask both speakers about this. There are, however, problems with taking self-reports of knowledge on face value. Indeed, there are a number of experiments in cognitive science showing that there is a gap between what people feel they know about something and what they are able to actually explain about it. For instance, people tend to overrate how well they understand a text (Glenberg & Epstein, 1985), how a device or a natural phenomenon works (Rozenblit & Keil, 2002), or even everyday objects (Sloman & Fernbach, 2017). Another series of studies has shown that people often fail to distinguish between their own knowledge and the knowledge that other people have (Sloman & Rabb, 2016). They also fail to appreciate that their perceived quality of a categorical explanation depends on a sense of how much the word is entrenched in the community, regardless of how much information they have about the word (Hemmatian & Sloman, 2018).

Following in the footsteps of several studies that empirically investigated the emergence of signaling conventions (Cochran & Barrett, 2021; Bruner *et al.*, 2018), in the present study we set out to determine whether people's self-perception of understanding only depends on their own knowledge about the word's reference, or if it is affected by the division of linguistic labor. To this end, we have developed an experiment in which the conditions for the emergence of division in a cognitive task, mediated by language, are instantiated in an image classification task. In the experiment, participants received points for correctly labeling dogs. Accurate labeling required participants to draw distinctions between two pairs of highly confusable pairs of breeds: Norwich vs Cairn Terrier, and Irish Wolfhound vs Scottish Deerhound. Half of the participants were in a solo condition in which they had to label the breeds on their own. The other half of participants were in a paired condition in which one participant was trained to distinguish between the two terriers while the other participant was trained to distinguish between the two hounds. After training, paired participants could ask their partner whether a label was correctly assigned to a dog image that they were tasked with labeling. We are interested in how a participants' perceived understanding of the dog breeds was affected jointly by their own independent ability to label dogs of different breeds, and by their being embedded in a dyadic micro-community in which their paired partner either did or did not know how to label breeds. The results of our experiment show that people's self-rated understanding was influenced not only by their accuracy at recognizing the word's reference, but also by their perceived knowledge relative to others.

2 Method

2.1 Participants and procedure

Participants were 84 students from the Universidad del Rosario (40% male, 60% female; average age $\approx$ 20 years), who were invited by e-mail to participate in the subject pool of the Rosario Experimental and Behavioral Economics Lab of the Universidad del Rosario. Participants were tested in 6 sessions; the first three corresponded to the treatment in pairs and the last three to the individual treatment. For the sessions in pairs on which an odd number of participants were received, the last participant to arrive was included in the individual treatment. The numbers of participants per session were: 3, 15, 24, 21, 8, 13. The total number of participants in the treatment in pairs was 40 and in the individual treatment was 44. Before starting the session, each participant signed an informed consent document, which was previously approved by the Ethics Committee (social science room) of the Universidad del Rosario. Each participant carried out the test inside a sound and sight isolated workstation running a version of the experiment implemented in the *nodeGame* platform (Balietti, 2017).[2] The test consisted of the presentation of instructions, followed by a brief quiz to confirm that the participants had understood the instructions in general. These provided the participants with knowledge about the game and the payment structure. After reading the instructions and passing the quiz, the participants completed a short tutorial on the operation of the task, after which they performed 25 training rounds and 25 game rounds. At the end of the game rounds, participants were asked to rate their understanding of each category on a scale of 1 to 7 (1=little or no understanding, 4=moderate understanding, 7=deep and detailed understanding), using a questionnaire adapted from (Rozenblit & Keil, 2002).

For their participation in the experiment, participants received 10 Experimental Monetary Units (EMUs; each unit equivalent to approximately 0.3 USD). Additionally, participants could earn bonus money based on their performance during the experiment. In both treatments, the payment received by each participant was determined by randomly choosing two game rounds and observing the score obtained in each of them. A score of 1 was rewarded with 2 EMUs; a score of 2 with 4 EMUs; a score of 3 with 6 EMUs; a score of 4 with 9 EMUs; and a score of 5 with 13 EMUs, so each participant could earn up to 26 EMUs additional to the show up fee. Participants received their reward in cash immediately after the session.

2.2 Materials

Regardless of the treatment, each participant was randomly assigned with equal probability to either the group trained to classify two Hound breeds or the group trained to classify two Terrier breeds (see Figure 1). During the training rounds (see Figure 2), each participant had to learn by trial and error how to correctly classify five dogs that appeared on their screen. However, participants being trained to discriminate between the two Hound breeds were provided with a number of explicit characteristics that only Irish Wolfhounds have (e.g., their hair color and build); and participants being trained to discriminate between the two Terrier breeds were given some explicit clues about Norwich Terriers.

[2] The reader can consult the experimental test at [url removed to ensure anonymity].

Cairn Terrier | Norwich Terrier | Irish Wolfhound | Scottish Deerhound

Figure 1: Examples of the four breeds of dogs used in the experiment. It requires some familiarity with the exemplars to be able to differentiate between the two breeds of Terriers and the two breeds of Hounds.

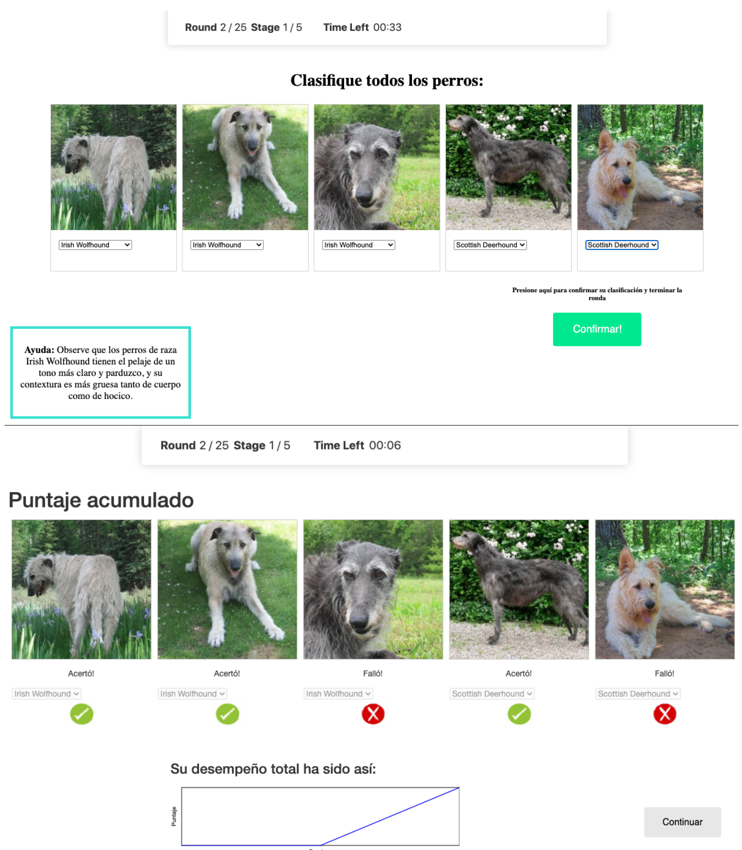

Figure 2: Example of the screen during the training rounds. Top: Classification step; Bottom: Score and feedback step.

At the end of the round, participants were told whether they had correctly classified each dog, but they were not provided with the correct label (see Figure 2 bottom panel). In the treatment in pairs, members of each pair were trained on complementary breeds (i.e., one participant was trained on Hounds and the other on Terriers). This information was explicitly given to the participants.

After the training rounds, players went on to the game rounds (see Figure 3). There the task changed slightly, as the number of categories available to classify the five dogs increased from two to four. The dogs on screen, in turn, were also mixed (i.e., there were both Hounds and Terriers). The task at this stage was the same – to correctly classify each of the five dogs into one of the available categories. Both paired players had exactly the same dogs on their screens during each round. In addition, for paired players, communication was enabled through a button with which each participant could ask their partner if a certain dog belonged to a specific category. The player could ask "Is this dog . . . ?" and fill in the blank with the label of one of the four breeds. To that question, the partner could only answer "yes" or "no" (or could simply ignore the message).

At the end of each round, both in the training and in the game rounds, each participant received feedback on their performance during the round.

3 Results

First,[3] we verified that participants reached a level of expertise in their respective *expert dog* breeds.[4] In Figure 4 on the left, it can be seen that after 25 rounds of training, participants achieved more than 90% accuracy, and this result is comparable in both treatments. Accuracy for expert dog classification was maintained during game rounds, which can be seen in Figure 4 in the center. In Figure 4 to the right we observe the development of accuracy for the novice dogs, in which paired novices did a slightly better job than solo novices, although this difference is not statistically significant.

We also verified that paired participants communicated in the expected way. First, we verified that novices asked experts about their novice dogs. Figure 5 on the top left shows the proportion of dogs of each breed on which the novices

[3]The reader can consult the data as well as the data analysis at [url removed to ensure anonymity].

[4]We will use the term "a subject's expert dogs" to refer to the breeds of dogs on which the subject received training. Similarly, the term "a subject's novice dogs" refers to breeds over which the subject did not receive training.

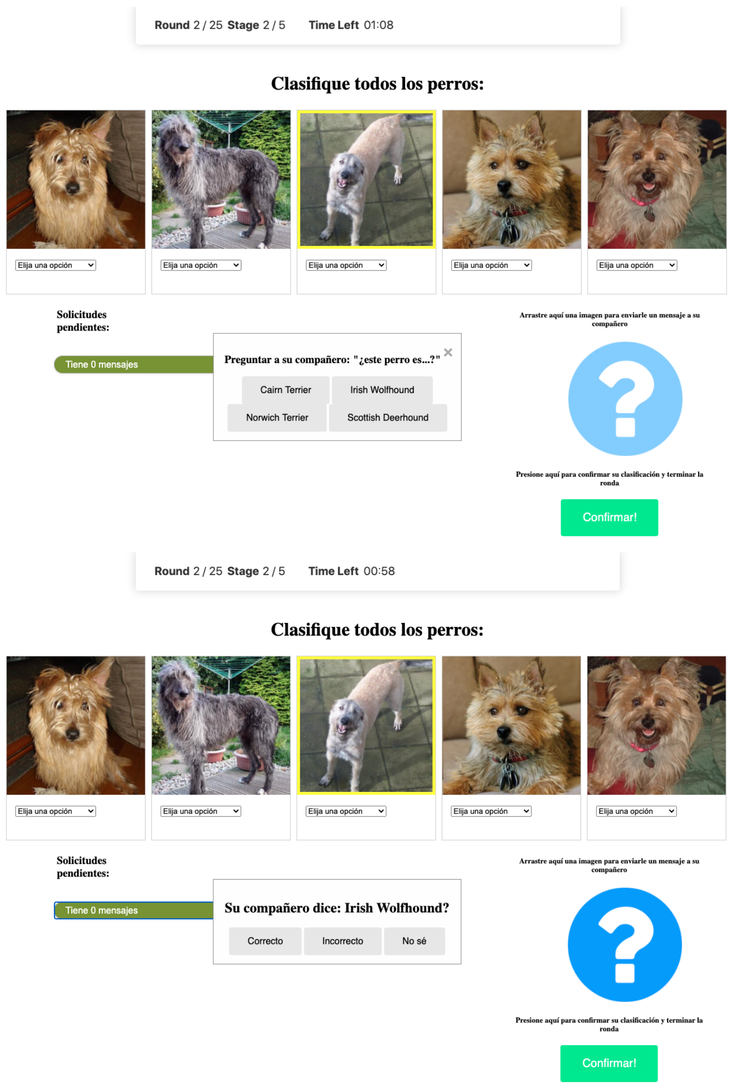

Figure 3: Game rounds. Top: Player 1 asks their partner about the dog category in the center image; Bottom: Player 2 receives the question from Player 1.

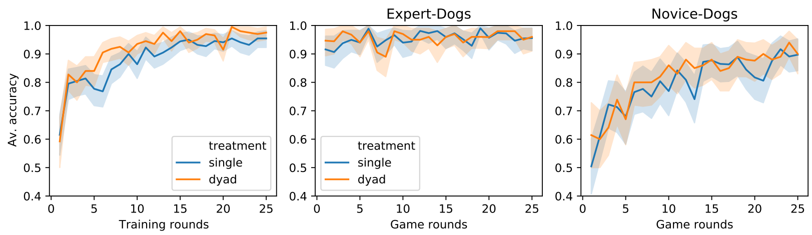

Figure 4: Accuracy in classification. Left: Training rounds. Center and right: Frequency of success in the classification of expert dogs (center) and novice dogs (right). The shaded region corresponds to a 95% confidence interval.

made a query to the expert. It can be seen that the novices asked their partners about 40% of their novice dogs, in contrast to less than 1% about their expert dogs. In contrast to the modestly inquisitive performance of novices, the data suggest that experts were much more active in answering queries. Figure 5 at the top right shows that over 90% of messages were answered, and of these over 80% were answered correctly. Although the trend in the number of messages sent decreased somewhat as the rounds progressed – probably because novices were learning to become experts with the feedback from experts –, novices continued to send a significant number of messages per round, as can be seen in Figure 5 at the bottom left. It is important to note that novices who tended to send more queries were more likely to be accurate, as can be seen in Figure 5 at the bottom right (corr. coef. ≈ 0.328).

Now we consider the results of the reports on category understanding, according to the questionnaire that participants had to answer at the end of the experiment. In Figure 6 above we can see the report distributions. Focusing first on experts, we observe that paired participants reported on average a greater understanding of their expert dogs than solo participants (p-value $<$ 0.001). This is interesting because the difference in experts' performance on both treatments is not statistically significant, as we had seen before (see Figure 4 center). Observe that accuracy does not make a significant contribution to the report for dyadic experts (see Figure 5 bottom right), as compared to a significant influence in the solo condition (see Figure 5 bottom left). In fact, the correlation coefficient between report and accuracy for solo participants is around 0.3, whereas it is only 0.15 for dyadic participants and the regression coefficient is not significant when regressing report with respect to accuracy (p-

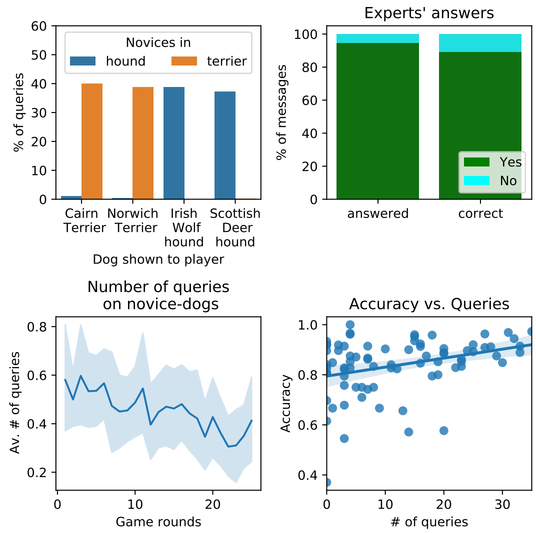

Figure 5: Use of labels. Top left: Percentage of dogs on which players made a query, separated by each of the breeds and by each type of novice dog. Top right: Percentage of messages answered by experts and percentage of correctness of the answer. Bottom left: Average number of messages sent for each round of play. The shaded region corresponds to a 95% confidence interval. Bottom right: Accuracy vs. player's query rate about their novice dogs. Each data point corresponds to a novice dog breed of each player in the pair treatment (N=160).

value $\approx$ .183). The increase in the report of understanding in the case of paired participants seems to be due to the fact that they were embedded in a dyadic micro-community. This idea receives additional support when considering the contribution to the report made by the expert's response rate to the novice's questions. In fact, when regressing the report with respect to accuracy and response rate, we observe that both coefficients are positive, although the accuracy coefficient is not significant (see Figure 6, bottom right; p-value $\approx$ 0.23) and the second one is marginal (see Figure 6, top right, p-value $\approx$ 0.056).[5] This result suggests that even after discounting the contribution of accuracy to the report, the response rate (albeit only slightly significant) is positively correlated with the report.

In Figure 6 top left we can see the distribution of the reports of understanding by *novices* in both treatments. It can be seen that, although the averages are

[5] To perform the aforementioned regression, we left out an outlier observation. One of the participants reported a score of 1 with respect to one of their expert dog breeds, after performing badly in the classification task.

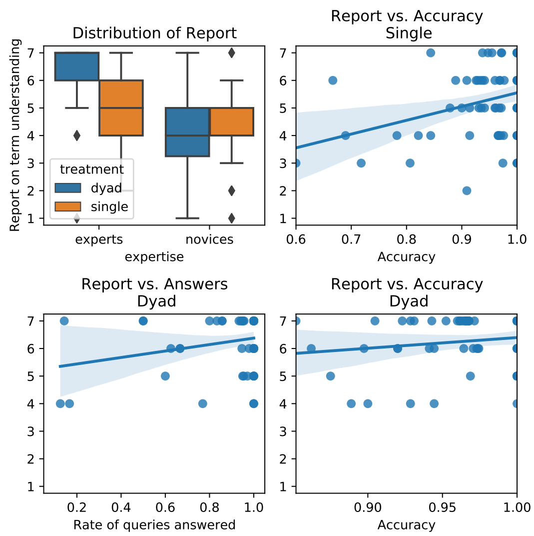

Figure 6: Reports on category understanding. Top left: Distribution of ratings in the reports, grouped according to treatment and separated on the x-axis by expert and novice dogs. Bottom: Influence of accuracy on reports in the case of experts for solo (left) and paired (right) participants. Each data point is a breed of expert dog for each player (N=80 left; N=88 right). Top right: Influence on the report of response rate to questions asked by the novice (N=88).

not so different, the variance of reports for paired novices is much greater than that of solo novices. Explaining this variance for the latter does not seem very difficult, since here the report is highly correlated with accuracy (see Figure 7, top left; corr. coef. ≈ 0.26). But when we focus on paired novices we find very interesting trends. First, when we try to explain the variation of the report with respect to accuracy, we find that the correlation is very low (see Figure 7 on the right; corr. coef. ≈ 0.07). Additionally, something interesting appears when we regress the report with respect to accuracy, the number of queries, and the rate of response to these queries by the expert. The regression coefficients for the first two are not significant (p-value accuracy ≈ 0.51; p-value number of queries ≈ 0.85) while the coefficient for the third is negative and significant (see Figure 7, bottom right; p-value ≈ 0.033). This means that as the expert responds to fewer of the novice's queries, the report of understanding by the novice increases. Probably this is because an uncooperative partner leaves no choice but to learn the correct ratings by oneself, which is reflected in a higher report of understanding at the end of the experiment.

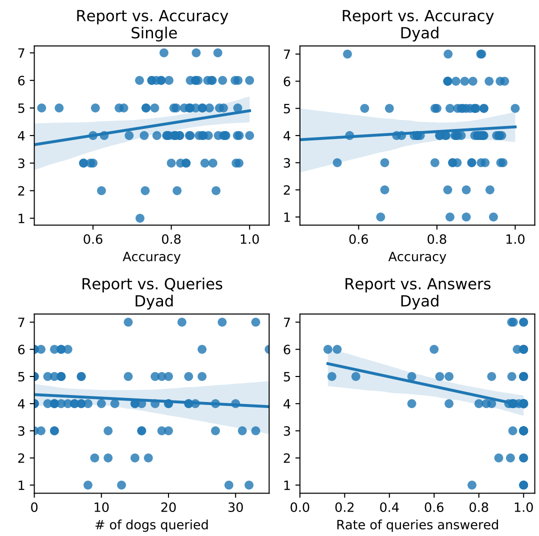

Figure 7: Reports on category understanding by novices. Top: report vs. classification accuracy for participants in the individual treatment (left) and in the treatment in pairs (right). Bottom: In the case of treatment in pairs, the report vs. the number of queries made by the participant about their novice dogs (left) and the report vs. the expert's response rate to these queries (right). Each data point is a breed of expert dog for each player (N=88 on the top left; N=80 on the rest).

4 Discussion

Our experimental paradigm succeeded in capturing several aspects of DLL found in the broader social world. We succeeded in creating experts through training. Training caused participants to become experts at a particular kind of dog (Terrier vs Hound). Discriminating breeds within each kind of dog (e.g. Terrier vs Hound) was difficult, but gradually improved over the course of 25 training rounds. This expertise carried over to a testing phase, leading to improved performance on the trained "expert" dog breeds compared with intermingled novel, untrained breeds. Furthermore, the differential expertises attained by members within a dyad was used by the members themselves. The evidence for this was that novice participants deferred to experts by querying them about unfamiliar dog breeds about 40% of the time. Experts were even more likely to actively engage in their role as such, responding to novices' queries about 90% of the time. Furthermore, novices who make more queries of their expert partners eventually reached higher levels of accuracy. Thus, we were able to create cognitive specializations among the players in dyads, and the players themselves recognized their acquired knowledge differences, and used these differences to produce improved collective performance in a labeling task.

At the broadest level, people's self-rated understanding was strongly influenced by their social context - whether they were labeling dogs on their own or with possible help from a peer who was given different training. However, the direction of this influence was surprising by some accounts of DLL. According to an account in which a person believes that they know more because they are connected to other people who know things which they do not, we would have predicted self-rated understanding to be higher for novices in dyads than solitary novices. In fact, that difference was not significant. An additional piece of evidence against this account is that novices rate their own understanding as lower when their partner expert answers more of their queries. A stronger connection to a trustworthy expert partner would be expected to increase the novice's sense of understanding if the novice gave themselves "partial credit" for knowing what their partner knows. Instead, in our paradigm, it seems that the novice feels that they understand the dog categorizations more when they know them independently, without needing to defer to others. This result differs from Sloman & Rabb (2016) result that being connected to a knowledgeable other increases one's own sense of knowing. A likely reason for this difference is that in our paradigm participants are very well aware when they don't know the difference between, say, a Cairn Terrier and a Norwich Terrier. When they resort to asking their expert partner, it is because they realize that they are uncertain themselves. That is, a novice feeling forced to ask their expert partner about a dog breed might be the equivalent to Rozenblit & Keil (2002) showing participants that they do not know how a mechanism works by asking them about the mechanism. Only after people have tried to answer a question themselves do they become aware that they are not as knowledgeable as they originally thought. In our experiment, if the novice's expert partner does not respond to their queries, then the novice takes it upon themselves to learn the distinction between the breeds, which leads to their subsequent impression that they do have an independent understanding.

Instead of finding higher self-rated understandings for dyadic than solitary novices, we found that self-rated understanding is higher for experts in dyads than solitary experts. This is apparently not due to experts believing their partner knows more about the expert dogs than do the experts. Experts very rarely made queries to their partners about breeds for which they were experts.

Instead, a plausible hypothesis is that when the novice partner queries the expert for advice, then the expert's confidence in their own knowledge increases. This is consistent with recent empirical results showing that the act of giving advice increases one's confidence Eskreis-Winkler *et al.* (2019). In fact, in one study, giving advice raised self-reported confidence more than receiving advice did Eskreis-Winkler *et al.* (2018). This latter result is consistent with our finding that experts giving advice increased their self-reported understanding more than did novices receiving advice from experts. The increase in experts' self-reported understanding when they are put in dyads was not, alas, accompanied by actual improvements in their accuracy. Experts in the dyad and individual conditions did not differ in their actual accuracies. Accordingly, the experts' metacognitive judgments (i.e. a judgment about one's own cognition) were biased by their social context. This biasing effect of social context is also suggested by the result that for single experts, their accuracy was significantly related to their self-perceived understanding, but that is no longer true when the experts were in dyads. Self-perceived understanding of experts in dyads is strongly influenced by variables other than accuracy. One salient variable is how often the expert was queried by their partner. A broad sociological pattern is that an expert's status is often related to how often they are queried for advice Grundmann (2017), and so it is plausible that experts used queries from others as informative about their own level of knowledge, thereby weakening the connection between self-reported understanding and actual accuracy. This weakened connection due to dyadic context is also found for our novices. Thus, our results speak to other factors besides accuracy looming larger when people judge their knowledge in dyads compared to when they are isolated. A good candidate for one such factor is that one's perceived knowledge *relative to others* looms large in social contexts.

Returning to the notion of Division of Linguistic Labor that motivated our experiment design, our experiment makes clear conflicting predictions derived from philosophical psychology versus social psychology. The first prediction is that, contrary to more traditional philosophical accounts that focus only on the individual, an individual's perception of understanding of a word might be fixed by what other people within their community know rather than by the individual themself. The second prediction is that people often assess their own skills and attributes by comparing themselves to others (Wood, 1996). Our evidence suggests that, albeit the individual's skills are not the whole story, the second effect is empirically larger than the first. People may have an intuition that corresponds to 'world-involving understanding', manifested by increased feelings of understanding word meanings when they are interacting with an expert who can competently match the word to perceptual inputs than when they are isolated. However, if so, this intuition is eclipsed by a stronger intuition that the expert understands a lot when they understand more than their paired novice. This conflict is likely to arise in many real-world social situations. If we want people to appreciate that they are more capable and effectively smarter when they are part of smart communities (Goldstone & Theiner, 2017), then strategic steps may need to be taken to block the natural tendency for people to compare their own abilities to others. Harnessing the potential of collective intelligence will likely involve people realizing that what matters most for solving societal problems is not how much I know compared to you, but how much we know together.

References

Briscoe, Robert Eamon. 2006. Individualism, Externalism and Idiolectical Meaning. *Synthese*, **152**(1), 95–128.

Bruner, Justin, O'Connor, Cailin, Rubin, Hannah, & Huttegger, Simon M. 2018. David Lewis in the lab: experimental results on the emergence of meaning. *Synthese*, **195**, 603–621.

Cochran, Calvin T., & Barrett, Jeffrey A. 2021. How signaling conventions are established. *Synthese*.

Eskreis-Winkler, L., Fishbach, A., & Duckworth, A.L. 2018. Dear Abby: Should I give advice or receive it? *Psychological Science*, **29**, 1797–1806.

Eskreis-Winkler, L., Milkman, K.L., Gromet, D.M., & Duckworth, A.L. 2019. A large-scale field experiment shows giving advice improves academic outcomes for the advisor. *Proceedings of the National Academy of Sciences*, **116**, 14808–14810.

Evans, Gareth. 1982. *The Varieties of Reference*. Oxford University Press. Edited by J. McDowell.

Glenberg, A.M., & Epstein, W. 1985. Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory & Cognition*, **11**, 702–718.

Goldstone, R. L., & Theiner, G. 2017. The multiple, interacting levels of cognitive systems (MILCS) perspective on group cognition. *Philosophical Psychology*, **30**(3), 334–368.

Grundmann, R. 2017. The problem of expertise in knowledge societies. *Minerva*, **55**, 25–48.

Hemmatian, Babak, & Sloman, Steven. 2018. Community Appeal: Explanation Without Information. *Journal of Experimental Psychology: General*, **147**(11), 1677–1712.

Kallestrup, Jesper. 2012. *Semantic Externalism.* Routledge.

Putnam, Hillary. 1975. The meaning of 'meaning'. *Pages 131–193 of: Minnesota Studies in the Philosophy of Science*, vol. 7. Minneapolis: University of Minnesota Press.

Rozenblit, L., & Keil, F.C. 2002. The misunderstood limits of folk science: an illusion of explanatory depth. *Cognitive Science*, **26**(5), 521—-562.

Sloman, Steven, & Fernbach, Philip. 2017. *The Knowledge Illusion: Why we never think alone.* Riverhead.

Sloman, Steven, & Rabb, N. 2016. Your understanding is my understanding: Evidence for a community of knowledge. *Psychological Science*, **27**, 1451–1460.

Talmage, Catherine J. L. 1998. Is there a division of linguistic labour? *Philosophia*, **26**(3-4), 421–434.

Wagner, Henri. 2020. The significance of the division of linguistic labor. *The Monist*, **103**, 381–390.

Ware, Robert. 1978. The Division of Linguistic Labor and Speaker Competence. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, **34**(1), 37–61.

Wood, Joanne V. 1996. What is Social Comparison and How Should We Study it? *Personality and Social Psychology Bulletin*, **22**(5), 520–537.