

The Role of Explanation in the Evolution of Technology

FRANÇOIS OSIURAK
Université de Lyon and Institut
Universitaire de France

GIOVANNI FEDERICO
Suor Orsola Benincasa University

CHLOÉ BRYCHE
Université de Lyon

ALEXANDRE BLUET
Université de Lyon and Karolinska
Institutet

MAXIMILIEN METAIREAU
Université de Lyon

ÉLÉONORE TRAN
Université de Lyon

JOËL BROGNIART
Université de Lyon

NICOLAS CLAUDIÈRE
Aix-Marseille Université

One way to pass on our own causal understanding about how a physical system works is to provide a verbal explanation about it, raising the question of the role of explanations in the transmission and evolution of technology. Here we address this question by conducting new analyses on the data collected in 3 previously published microsociey studies, in which participants organized into transmission chains had to improve a technology. In 3 of the 6 conditions reported, the participants were also asked to generate verbal explanations about how the technology works. In 1 of these conditions, these explanations were transmitted to the next participant. Our results indicated that the participants improved the technology and gained a better understanding of it over generations, although they were unable to elaborate on its hidden causes using scientific language. When the explanations were analyzed as instructions, we did find an increase in the quality of the explanations over generations, which was linked to participants' causal understanding. However, this increase had no direct impact on the transmission of the technical content. These findings open interesting avenues for future research on the role of explanation in the evolution of technology and raise the question of whether (explanation-based) teaching is an opportunity for the teacher to increase their own causal understanding.

KEYWORDS: causal understanding, cultural evolution, social learning, teaching

Recently, NASA's Mars *Curiosity* rover filmed a strange doorlike rock formation, which generated excitement over the Internet. In fact, what is really exciting is that humans are able to send a robot about

60 million kilometers away that can climb hills, take photographs and videos, and make analyses and decisions in an autonomous fashion, and with which we can communicate. This rover reflects a series of

American Journal of Psychology

Fall 2024, Vol. 137, No. 3 pp. 285–320 • © 2024 by the Board of Trustees of the University of Illinois

technical innovations that have been combined over time and that go far beyond what a single person could invent alone in their lifetime, illustrating the phenomenon known as cumulative technological culture (CTC) (Boyd & Richerson, 1995; Derex, 2022; Mesoudi & Thornton, 2018).

Like any other cultural phenomenon, CTC rests on social learning, which enables individuals to acquire technical knowledge through other individuals in a much easier and faster way than through asocial learning (Boyd, Richerson, & Henrich, 2011; Heyes, 1994; Osiurak & Reynaud, 2020; Tennie, Call, & Tomasello, 2009; Tomasello, Kruger, & Ratner, 1993; Whiten, 2021). All theories agree on this point. What is still debated is the cognitive skills that are needed for CTC to emerge. Recent perspectives suggest that scientists so far have tended to overemphasize the social dimension of CTC (Heyes, 2021a, 2021b; Osiurak & Reynaud, 2020; Perry et al., 2021; Singh et al., 2021; Vale et al., 2021; Whiten, 2022), minimizing the potential role of nonsocial cognitive skills such as causal understanding. Causal understanding—also called technical reasoning when applied to the technological dimension (Osiurak, Claidière, & Federico, 2023; Osiurak & Reynaud, 2020)—is the ability to reason about cause–effect relationships, by deriving predictions for interventions after learning by observing events (Gopnik et al., 2004; Gopnik & Schulz, 2004; Pearl, 2000; Waldmann & Hagmayer, 2005). Thus, some have assumed that the reconstruction nature of causal understanding is essential not only for innovation but also for high-fidelity transmission, the two key drivers of CTC (Osiurak et al., 2022, 2023).

How is causal information passed on between people during social learning? There are many ways that are often combined and complementary, such as the observation of someone else's behavior, a demonstration of what to do, or the provision of explanations (Morgan et al., 2015). The role of explanations in the social transmission of technology has received little attention so far. One reason is that explanations are not necessary for the emergence of CTC and are often considered an elaborate form of social learning that is typically human (since it relies on language). Nevertheless, humans, at least in more recent times, tend to provide explanations in many contexts of social transmission (e.g., formal or informal apprenticeship, school, user instructions). Therefore, the role

played by explanation in the social transmission of technology deserves attention.

Explanation differs from teaching, which can be broadly defined as behavior that facilitates learning in others (Caro & Hauser, 1992; Kline, 2015). Thus, someone can generate explanations to teach someone or to try to complete their own understanding (Keil, 2006), and someone can teach by generating explanations (i.e., verbal teaching) or not (e.g., gestural teaching or demonstration) (Gardenfors & Högberg, 2017). People create explanations for everything, regardless of their level of expertise in the domain (Keil, 2006; Keil, Rozenblit, & Mills, 2004). Thus, laypeople can explain why the harvest will be bad this year, and experts why the climate is changing, and vice versa. This type of explanation refers to theoretical explanation (i.e., scientific language), which reflects an attempt to directly communicate the understanding of a causal effect to another person. There is another type of explanation, practical explanation (i.e., instructions), which is a linguistic description of how to do something effectively without providing any specific information about the underlying causal effects. This is very close to some forms of nonverbal teaching, such as when the model orients the learner's attention toward salient parts of the demonstration by adding gestures (e.g., pointing) or modifying the target behavior (e.g., repeating or slowing down some relevant actions). In this case, elements of understanding can be transmitted indirectly by placing the learner in good conditions to acquire them. However, there is no guarantee that these elements of understanding are transmitted (see Jagiello, Heyes, & Whitehouse, 2022). This is also true for theoretical explanations.

The goal of this article is to elucidate the role of explanation in the social transmission of technology. As technologies can be passed on through several modes of transmission that are often combined, we will follow an “all else being equal” approach by holding all other factors equal and testing the role of explanation. Our main question will be addressed by answering two secondary questions. Evidence indicates that the improvement of a technology over generations is accompanied by an increase in its understanding (Osiurak et al., 2021, 2022; but see Derex, Bonnefon, Boyd, & Mesoudi, 2019). Thus, the first question (Q1) is whether this is also true

for explanations. Is the improvement of a technology over generations accompanied by an increase in the quality of the explanations about how it works? A first hypothesis is that the quality of explanations increases over generations in parallel to the improvement of the technology and its understanding. If so, this suggests that explanations can facilitate mind-to-mind transmission by helping people transmit overtly what they understand covertly. A second hypothesis is that there is no increase in the quality of explanations. If so, this suggests that explanations, despite their intrinsic communicative nature, are not necessary for the transmission of technical content. A third hypothesis is that the link between the quality of explanations and the progressive improvement of technology depends on the nature of the explanation (i.e., practical vs. theoretical). The second question (Q2) is whether explanation transmission reinforces the presence of traditions until it provokes a detrimental canalizing effect. Indeed, some have suggested that explanation transmission might have a detrimental canalizing effect by shaping the exploration of parameter space (Drex et al., 2019). Thus,

inheriting an explanation about one dimension of a physical system might reduce the exploration of the other dimensions of the system.

To address the two questions, we conducted new analyses on the data collected in three previously published microsociey studies ($n = 420$; six experimental conditions) (Drex et al., 2019; Osiurak et al., 2021, 2022), which used the same physical system (i.e., a wheel that descended a 1-m-long inclined track) (Figure 1a), thus modifying its moment of inertia and the position of its center of mass. The general procedure of the three studies was the same (Figure 1b). The participants had five trials to improve their wheel and were presented with an understanding test after the five trials. For each condition, the participants were organized into 14 chains of five participants (or generations) and could transmit information (depending on the condition) about their last two trials to the next participants in the chain (except those in the last generation) (Figure 1c). In the configurations condition, the configura-

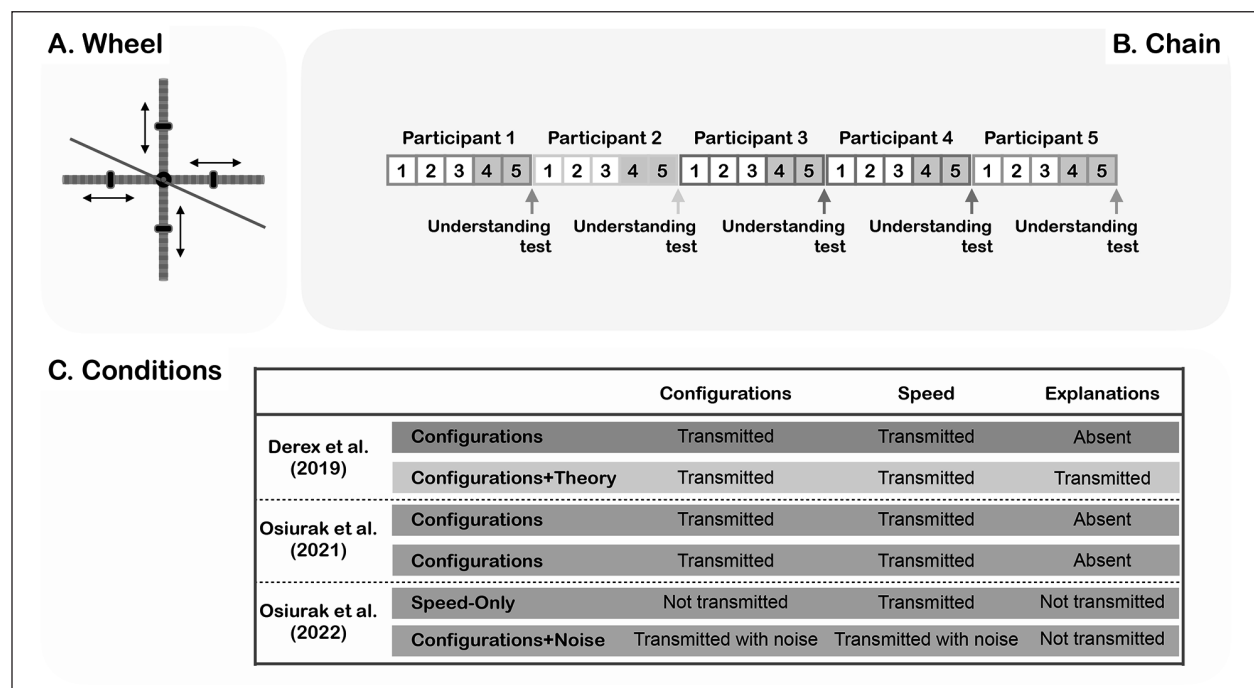


FIGURE 1. Overview of the experimental paradigm. (a) The task was to minimize the time it took a wheel to travel down a 1-m-long track by moving four weights to any of 12 discrete positions along each spoke. (b) The participants had five trials to improve the wheel system and were presented with an understanding test after the five trials. They were organized into chains of five individuals (or generations) and could transmit information about their last two trials to the next participants in the chain (except those in the last generation), which depended on the condition (c).

tions and the associated speed were transmitted. In the configurations+theory condition, the configurations and the associated speed were transmitted along with a 340-character explanation of how the wheel system works. In the configurations+noise condition, noise was added to the configurations and the associated speed of the last two trials before transmission. In the speed-only condition, only the wheel speed was transmitted. In both the configurations+noise and speed-only conditions, the participants were also requested to generate a 340-character explanation of how the wheel system works. However, contrary to the configurations+theory condition, these explanations were not transmitted to the next participants in the chain.

STUDY

METHODS

Participants

In total, 420 participants took part in the three studies ($n = 140$ for each study with $n = 70$ for each condition; 70 women and 70 men in Derex et al. [2019], 92 women and 48 men in Osiurak et al. [2021], and 102 women and 38 men in Osiurak et al. [2022]). The participants were mostly students, who were nonselectively recruited (mean age = 20.5 years, $SD = 3.4$ years in Derex et al. [2019]; mean age = 21.4 years, $SD = 3.8$ years in Osiurak et al. [2021]; mean age = 20.8 years, $SD = 2.9$ years in Osiurak et al. [2022]). Informed consent was obtained from all participants, and the three studies were approved by three ethics committees (University of Exeter, Catholic University of Lille, and University of Lyon). Participants were paid in Derex et al. (2019) and were not in Osiurak et al. (2021, 2022).

Experimental Apparatus

The wheel system was originally developed by Derex et al. (2019). The system used in Osiurak et al. (2021, 2022) was not an exact reproduction; the main difference concerned the axis of the wheel, which was not made of the same materials and which generated differences in terms of wheel speed in the different studies (for more discussion about this aspect, see Osiurak et al., 2021). To make the wheel speeds comparable between the six conditions of the three studies, we normalized the wheel speeds by condition using a z score transformation with a mean of 167

m/hr and a standard deviation of 14 m/hr. The mean and the standard deviation chosen were based on the complete dataset ($n = 2,100$; i.e., 5 trials $\times$ 70 participants $\times$ 6 conditions) and corresponded to the mean and the SD computed from all the speeds of wheels that descended the track (without the wheels with a speed of 0 m/hr). Illustrations of the wheel system and more details about it are available at <https://osf.io/afwmr/> and <https://osf.io/m3d7q/>.

Procedure

The experiments took place in experimental rooms (around 20 to 30 min in duration). The participants sat at a table placed 2 m from the experimental apparatus and could use a computer to indicate their response in Derex et al. (2019). By contrast, Osiurak et al. (2021, 2022) used a paper-and-pencil version. The participants were instructed that their task was a collective task. The participants had five trials to optimize the speed of a wheel that traveled down a 1-m-long inclined track by moving four weights to any of 12 discrete positions along each spoke (from 1 to 12, with 1 being the closest to the center of the wheel and 12 being the farthest from the center of the wheel). The time it took the wheel to descend the track was recorded by a computer program. The information about the last two trials was available to the participants, and they had as much time as they needed to consult it. After three trials, the experimenter reminded them that the information about their last two trials would be transmitted to the next participant in the chain. The nature of information transmitted to the next participant in the chain varied according to the condition. After the building phase, the participants completed an understanding test and, in some conditions, were invited to write a brief theory (less than 340 characters long) about the functioning of the wheel system, which started with “The wheel covers the distance faster when. . .”

CONDITIONS.

Each study included two conditions (six conditions in total) with 14 chains of five participants in each (70 participants per condition). The conditions were a configurations condition and a configurations+theory condition in Derex et al. (2019), a two-configuration condition in Osiurak et al. (2021), and a speed-only condition and a configurations+noise condition in Osiurak et al. (2022). In the configuration conditions, the configurations and the associated speeds of the last two trials were transmitted to the next participant in the chain. In the configurations+theory condition,

the configurations and the associated speeds were also transmitted along with the 340-character-long explanation. In the speed-only condition, only the speed of the wheels of the last two trials was transmitted (no information about the configurations). In the configurations+noise condition, the experimenter introduced a random noise into the last two wheel configurations before transmission by moving the four weights six positions farther from or closer to the center of the wheel (the absolute sum of the modifications was 6). The configurations thus modified and their associated speeds were then transmitted to the next participant. Note that the 340-character-long explanation generated by the participants in the speed-only and configurations+noise conditions were not transmitted to the next participant in the chain.

UNDERSTANDING TEST.

During the understanding test, participants had to determine which of the wheels would cover the track faster. In the first version (Derex et al., 2019), the test was composed of 10 items (five inertia items and five center-of-mass items). There were two configurations for each item, and the participants could choose between three options (wheel 1, wheel 2, and no difference). The no difference option was always incorrect. The possible range of scores was 0–10 for the total score and 0–5 for each dimension (i.e., inertia or center of mass). Several versions were developed in Osiurak et al. (2021). Here we conducted our analyses on the test called “analogous test.” In the configurations condition of Experiment 1, there were 24 items, each with four options. The incremental difference between the four options was 5 m/hr (eight items), 10 m/hr (eight items), or 15 m/hr (eight items). Each item was scored from 1 (the wheel with the lowest speed) to 4 (the wheel with the highest speed). Thus, the possible range of scores was 24–96. There was no subscore for each dimension (inertia or center of mass). In the configurations condition of Experiment 2, the test also consisted of 24 four-option items, each scored from 1 to 4 (from the lowest to the highest speed). However, there were 12 inertia items and 12 center-of-mass items. Therefore, the possible range of scores was 24–96 for the total score and 12–48 for each dimension (inertia or center of mass). This version of the understanding test was also used for the speed-only and configurations+noise conditions in Osiurak et al. (2022). To make the understanding scores comparable between the six conditions of the three studies, we normalized the understanding scores by condition using a *z* score transformation with a mean of 100 and *SD* of 15,

as in standard IQ tests. The versions of the understanding test are available at <https://osf.io/afwmr/> and <https://osf.io/m3d7q/>.

CONTENT ANALYSES.

The 340-character-long explanations were generated by the participants in the configurations+theory condition (Derex et al., 2019) and in the speed-only and configurations+noise conditions (Osiurak et al., 2021, 2022), three conditions of 70 participants each (*n* = 210). Only 209 explanations were analyzed, one explanation being missing in the speed-only condition because of a technical error.

The first content analysis consisted of coding manually the information contained in the explanations according to two predefined categories, practical and theoretical. Practical information was any information about the position of the weights on the spokes (e.g., “the weights are close to the center,” “the top weight is at the highest position”). This kind of information can be viewed as instructions because it limits the need for the reader to make interpretations. By contrast, theoretical information was any piece of information that described the behavior of the wheel on the track via “scientific language” (e.g., “center of gravity,” “principle of inertia,” “to unbalance the wheel,” “at each revolution the wheel can lose energy”), requiring the reader to translate this information into actual positions of the weights. We coded each piece of information for each explanation into these two types of information (e.g., one explanation might have two practical and three theoretical pieces of information). This coding was performed by a first experimenter (F.O.), who was blind to condition, and repeated twice on the whole dataset at an interval of 1 week. The intrarater reliability was assessed via Cohen’s κ coefficient. The agreement was almost perfect (κ = .96). The coding was done again on the whole dataset by another experimenter (C.B.), more than 1 year after the initial coding performed by the first experimenter (F.O.). This second experimenter was also blind to condition. The interrater reliability was assessed via Cohen’s *k* coefficient, which revealed that the agreement was almost perfect (*k* = .84).

The second content analysis consisted of rating the explanations in terms of their quality as instructions. To do so, five judges (F.O., C.B., A.B., M.M., and E.T.), who were blind to condition, rated independently the explanations on two 0- to 3-point scales, one for the inertia dimension and the other for the center-of-mass dimension. The criteria used to rate each dimension are given in Table 1. The more

TABLE 1. Rating of the Explanations in Terms of Their Quality as Instructions

Criterion	
<i>Inertia dimension</i>	
3 points	The relation between the position of the weights and the wheel speed is correct. The explanation clearly states that the weights must be close to the center of the wheel. The explanation mentions that the weight must be globally balanced.
2 points	The relation between the position of the weights and the wheel speed is correct. The explanation suggests that the wheel is slower when its weights are far from its center without emphasizing that the weights must be close to the center.
1 point	The explanation mentions that the weight must be globally balanced.
0 point	None of the above information is mentioned.
<i>Center-of-mass dimension</i>	
3 points	The relation between the position of the weights and the wheel acceleration is correct. The explanation states that the top weight must be far from the center of the wheel. The other three weights are correctly positioned (i.e., generally close to the center of the wheel).
2 points	The relation between the position of the weights and the wheel acceleration is almost correct. The explanation states that the top weight must be far from the center of the wheel. The other three weights are not correctly positioned; notably, the front weight is too far from the center of the wheel relative to the top weight.
1 point	The relation between the position of the weights and the wheel acceleration is incorrect. The explanation mentions a relation between the wheel speed and an unbalanced repartition of the weights.
0 points	None of the above information is mentioned.

the instructions allowed participants to build with precision a wheel that obeyed the inertia and center-of-mass principles, the greater the instructions score. The possible range of scores was 0–30 for the total score and 0–15 for each dimension (inertia or center of mass). The interrater reliability was assessed via Kendall's τ coefficients. All coefficients were significant and positive (inertia dimension, from $\tau = 0.50$ to $\tau = 0.77$, all $p < .001$; center-of-mass dimension, from $\tau = 0.75$ to $\tau = 0.85$, all $p < .001$), indicating good interjudge agreement.

The third content analysis aimed to assess the relevance of the terms used in the explanations with respect to scientific definitions of the two main physical principles that governed the wheel system: inertia and center of mass. To do so, we collected from the internet 10 scientific definitions of inertia and 10 of center of mass. The only modification we made on these definitions, before importing them into the IRaMuTeQ software (Ratinaud, 2009) for textual analysis, was to add a connector “_” for the terms

center of mass (center_of_mass), center of gravity (center_of_gravity), and center of inertia (center_of_inertia) to distinguish these terms from the generic term “center.” The IRaMuTeQ software offered the possibility of lemmatizing the text by transforming the words and expressions into terms through the radicals of the words. This analysis allowed us to obtain the relative frequencies for the key terms used for each type of definition (inertia vs. center of mass) and to create a weighted score associated with each key term (the more frequent the key term, the higher its score). Three terms were removed because of their use in task instructions and during interactions with the participants (*position*, *speed*, and *center*). Finally, we coded the presence or absence of these key terms in the explanations provided by the participants, regardless of whether the key terms were used appropriately, and we computed for each explanation a scientific language score (sum of weighted scores of key terms). More information about the methods used here is available at <https://osf.io/yghxu/>.

SIMILARITY ANALYSIS.

For each condition separately, we computed two similarity scores, which were based on the configuration of the wheel that was fastest between the last two trials of each participant. For the within-chain similarity score, we calculated the sum of the distances between the weights on each spoke for all pairs of best configurations in the chain. This dissimilarity score could range from 0 to 44, and to obtain a similarity score we subtracted the dissimilarity score to 44. The procedure was the same for the between-chain similarity score except that we compared the best configuration of each participant with the 65 best configurations generated by the participants of the 13 other chains (for a similar procedure, see Osiurak, Cretel, et al., 2020, 2022; Reindl & Tennie, 2018). This similarity score was computed for each of the six experimental conditions ($n = 420$).

EXPLORATION.

The exploration score corresponded to the difference between the smallest and the greatest sum of positions of the four weights obtained during the five trials. The greater this difference, the greater the exploration. The center-of-mass exploration score corresponded to the surface of the convex envelope that contained the center-of-mass coordinates of the five wheels. The greater this surface, the greater the exploration. More details about these two exploration scores are provided in Osiurak et al. (2022).

STATISTICAL ANALYSES.

We used regression modeling in R (R Development Core Team, 2011), lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017), to fit linear mixed models (LMMs). The details of all the models fitted in the present study are given in Tables S2–S22 along with their results. Statistical significance was set at $p < .05$, and a bootstrapping method was used to estimate 95% CI.

RESULTS

Wheel Speed and Understanding

As shown in Figure 2a, an LMM analysis performed on the six experimental conditions indicated an increase over generations of both the wheel speed (generation estimate, 2.82 m/hr; 95% CI [2.12, 3.39]; Figures S1 and S2; Table S1) and the understanding score (generation estimate, 2.51; 95% CI [1.56, 3.49]; Figure S3; Table S2). A link was also found between both (wheel speed estimate, 0.69; 95% CI [0.56, 0.80]; Figure 2b and Figure S4; Table S3). These findings confirmed that the improvement of the wheel system was accompanied by an increase in its understanding.

Practical Versus Theoretical Information

We performed an exploratory analysis with the aim of characterizing the types of information (practi-

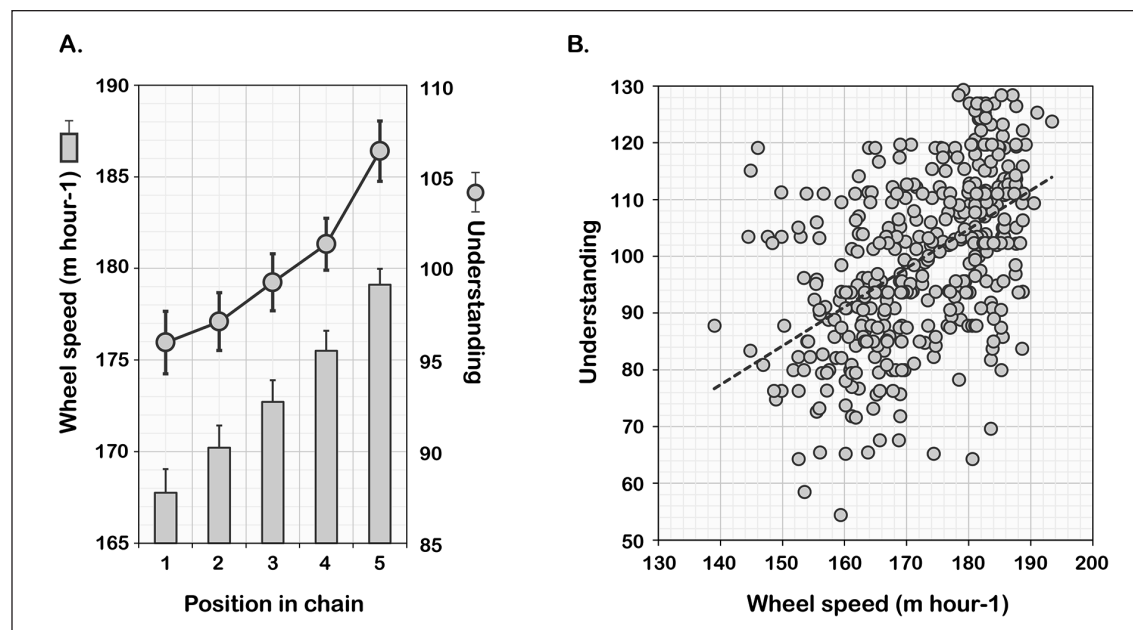


FIGURE 2. (a) Parallel improvement of the wheel system and its understanding and (b) link between the wheel speed and the understanding score. Error bars indicate standard errors

cal vs. theoretical) contained in the 340-character explanations generated by the participants in the configurations+theory, speed-only, and configurations+noise conditions, $n = 210$; mean, 199.58 characters; SE , 3.77 characters; range, 82–340 characters. An LMM analysis indicated that the total number of pieces of information (practical and theoretical) provided by the participants was positively linked to the number of characters used (number of characters estimate, 0.03; 95% CI [0.02, 0.03]; Figure S5; Table S4). The participants generated a greater number of pieces of practical information (mean, 2.28; SE , 0.08) than of theoretical information (mean, 1.88; SE , 0.12; theoretical information estimate, -0.40 ; 95% CI $[-0.03, -0.78]$; Figure S6; Table S5). Practical information was given earlier during the explanation (mean, 2.77th position of pieces of information given; SE , 0.09) than theoretical information

(mean, 4.14th position; SE , 0.13; practical information estimate, 1.08; 95% CI [0.76, 1.39]; Figure S7; Table S6). This first analysis led us to conduct two separate analyses about the quality of the explanations generated by the participants, one that focused on the quality of the instructions provided by the participants and the other on the scientific language they used.

Instructions (Q1)

LMM analyses revealed that the instructions score was positively linked to the number of pieces of practical information (practical information estimate, 0.98; 95% CI [0.59, 1.40]; Figure S8; Table S7) but not of theoretical information (theoretical information estimate, 0.01; 95% CI $[-0.25, 0.29]$; Figure S9; Table S8), which confirmed that the instructions score was based on practical information. As can be

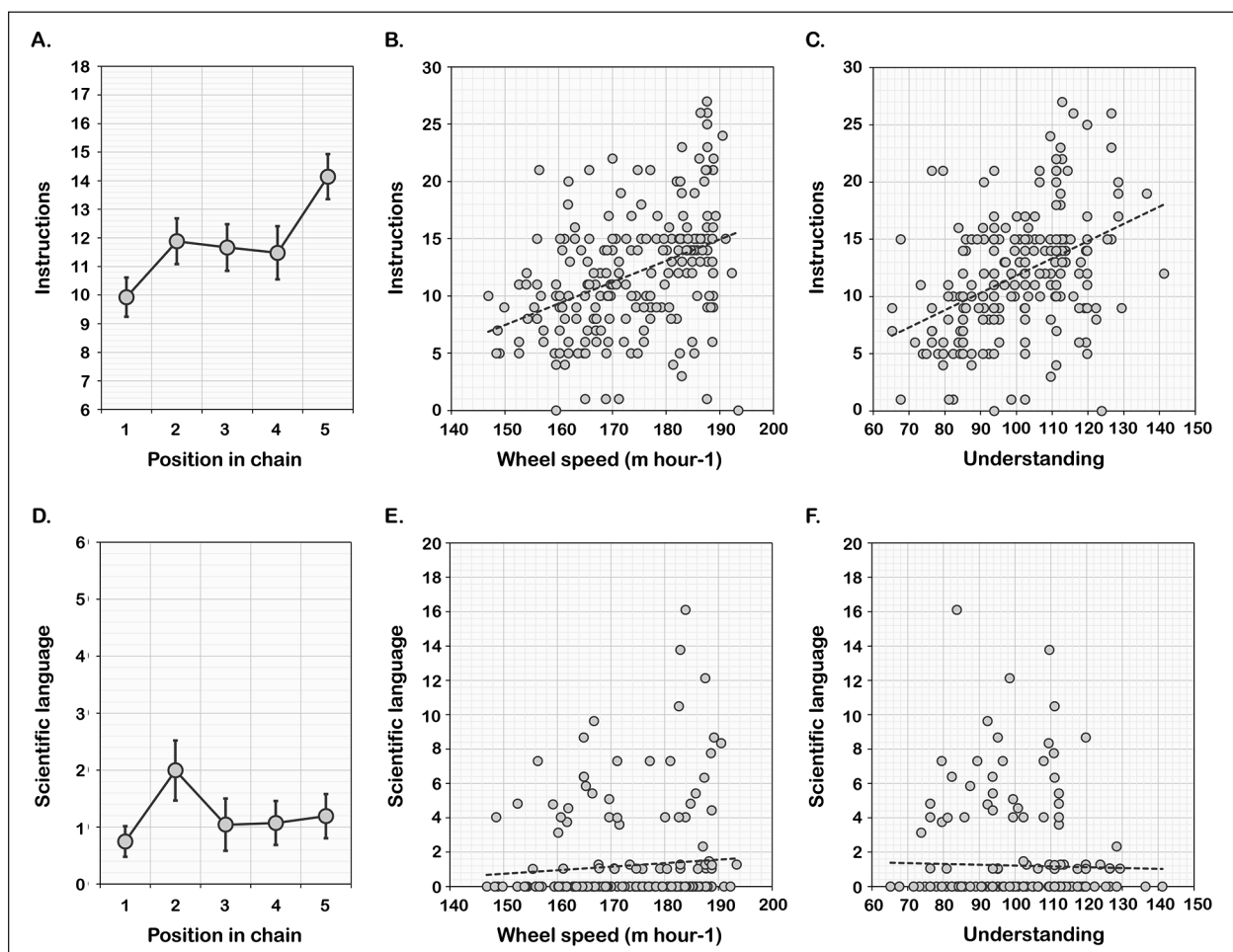


FIGURE 3. Increase in the instructions score (a) but not the scientific language score (d) over generations and links of the instructions score (b, c) or the scientific language score (e, f) with the wheel speed and the understanding score. Error bars indicate standard errors

seen in Figure 3a, the instructions score increased over generations (generation estimate, 0.79; 95% CI [0.30, 1.29]; Figure S10; Table S9). There was also a positive link between the instructions score and the wheel speed (wheel speed estimate, 0.19; 95% CI [0.13, 0.25]; Figure 3b and Figure S11; Table S10) and between the instructions score and the understanding score (understanding estimate, 0.15; 95% CI [0.10, 0.19]; Figure 3c and Figure S12; Table S11).

Scientific Language (Q1)

LMM analyses indicated that the scientific language score was positively linked to the number of pieces of theoretical information (theoretical information estimate, 0.65; 95% CI [0.54, 0.77]; Figure S13; Table S12) but not of practical information (practical information estimate, -0.23; 95% CI [-0.45, 0.01]; Figure S14; Table S13), which confirmed that the scientific language score was based on theoretical information. As shown in Figure 3d, the scientific language score did not increase over generations (generation estimate, -0.01; 95% CI [-0.25, 0.25]; Figure S15; Table S14). There was also no link between the scientific language score and the wheel speed (wheel speed estimate, 0.02; 95% CI [-0.01, 0.05]; Figure 3e and Figure S16; Table S15), nor between the scientific language score and the understanding score (understanding estimate, -0.01; 95% CI [-0.03, 0.02]; Figure 3f and Figure S17; Table S16).

Within-Chain and Between-Chain Similarity (Q2)

As can be observed in Figure 4 (see also Table S17), LMM analyses showed that within-chain similarity

scores were higher than between-chain similarity scores (within-chain estimate, 5.51; 95% CI [4.35, 6.71]). This within-chain versus between-chain similarity difference of 5.51 corresponded to that of the configurations condition of Derex et al. (2019), which was taken as the intercept. There were also several interaction effects, which indicated that this difference was higher in the configurations condition than in the two configurations conditions of Osiurak et al. (2021) (Configurations_Exp1*within-chain estimate, -2.63; 95% CI [-4.24, -1.05]; Configurations_Exp2*within-chain estimate, -1.75; 95% CI [-3.47, -0.28]) and in the speed-only and configurations+noise conditions of Osiurak et al. (2022) (Speed-Only*within-chain estimate, -5.87; 95% CI [-7.61, -4.18]; Configurations+Noise*within-chain estimate, -3.03; 95% CI [-4.84, -1.32]). Note that this interaction was so pronounced in the speed-only condition that the between-chain similarity score was even slightly greater than the within-chain similarity score in this condition. No interaction was reported with the configurations+theory condition (Configurations+Theory*within-chain estimate, -0.40; 95% CI [-2.16, 1.24]).

Exploration (Q2)

Finally, we explored whether the participants' exploration of either the inertia dimension or the center-of-mass dimension over the five trials was canalized by the explanation inherited from the predecessor. This so-called canalizing effect could occur only in the configurations+theory condition because it was the only condition in which the participants inher-

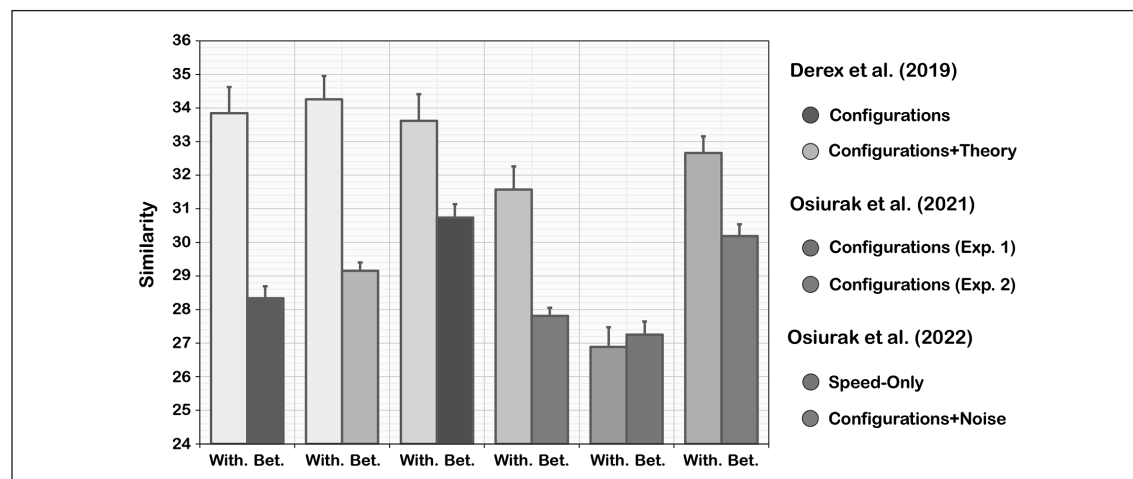


FIGURE 4. Within-chain and between-chain similarity as a function of condition. Error bars indicate standard errors

ited the explanations produced by their predecessor. We also report the results from the speed-only and configurations+noise conditions as control conditions (no explanation transmitted). We used the inertia and center-of-mass instructions scores to reflect the quality of the explanations inherited. The prediction was that the greater the quality of the instructions inherited on one dimension (e.g., inertia), the more the participants explored that dimension (i.e., iner-

tia) and the less they explored the other dimension (i.e., center of mass). As shown in Figure 5 (see also Tables S18–S21), this prediction was not confirmed. The main result was that participants in the speed-only condition explored the inertia dimension more than participants from the other conditions (Figure 5a: speed-only condition estimate, 12.42; 95% CI [5.40, 19.22]; Figure 5b: speed-only condition estimate, 7.39; 95% CI [0.89, 13.84]).

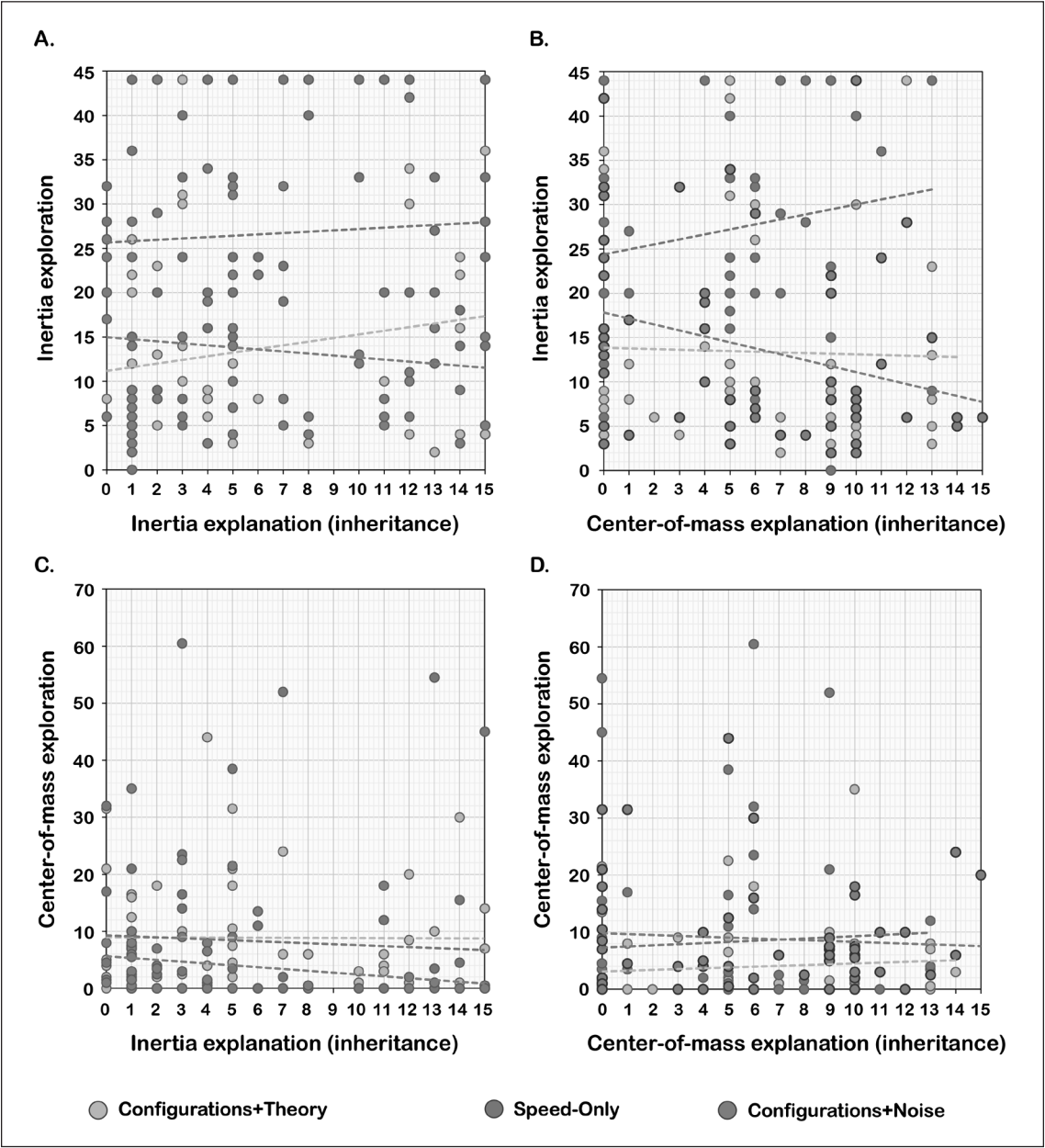


FIGURE 5. Links of inertia (a, b) or center-of-mass (c, d) exploration scores with inertia or center-of-mass explanation scores inherited

DISCUSSION

Evidence indicates that the improvement of a technology over generations is accompanied by an increase in its understanding. Our first question (Q1) was whether this is also true for the quality of explanations. The answer is yes and no. When explanations were analyzed as instructions, we did find an increase in their quality over generations, which was linked to the improvement of the wheel system and its understanding. This suggests that the quality of instructions could reliably reflect the participant's current state of understanding. Accordingly, we may expect that instructions favor the transmission of technical content between individuals. Yet nothing in the present data suggests that the participants who inherited explanations (those in the configurations+theory condition) outperformed the participants who did not inherit any explanation. This conclusion must nevertheless be tempered. Indeed, even if instructions could have helped some participants orient their attention toward salient features of the wheel system, their redundancy with the information contained within the configurations transmitted could have masked the potential benefits of inheriting them. Thus, the interest of instructions might be limited when participants can observe the technology itself—as long as it is not too opaque, as in the experiments reported here—or the actions performed by the model. Conversely, when the technology becomes opaque or neither the technology nor the materials needed to build it are available, using instructions might become the most effective—if not the only—way of transmitting technical content (Osiurak, De Oliveira, Navarro, & Reynaud, 2020). We did not report such conditions, which could offer interesting insights into the role of instructions in the transmission of technical content.

The participants did not only provide instructions to explain how the wheel system worked. They also used scientific language. In this case, and contrary to instructions, we did not find any increase in the quality of the scientific language over generations. This led us to also answer no to our first question (Q1). A potential methodological limitation concerns our way of creating the scientific language score, which took into consideration whether the terms were frequently used in scientific definitions, regardless of whether they were used appropriately.

Nonetheless, if the participants used scientific terms appropriately, the scientific language score should have been linked with the wheel speed or the understanding score, but it was not. This finding is consistent with earlier studies that have shown the difficulties people—including high school physics teachers—face in elaborating on general physical principles, whereas judgments are generally better in concrete contexts or when participants view ongoing events (Kubricht, Holyoak, & Lu, 2017; McCloskey, 1983; McCloskey, Caramazza, & Green, 1980; McCloskey, Washburn, & Felch, 1983). As stressed by Proffitt, Kaiser, and Whelan (1990), who asked participants to verbally answer questions on wheel dynamics, “after observing the wheels roll down the ramp, the subjects all came to the realization that mass distribution was the only influential variable. In general, their accounts of why mass distribution was influential simply described the particular wheels employed, and failed to elucidate general principles” (p. 350). In other words, it is one thing to possess and use semantic knowledge to explain a phenomenon and another to understand it, as demonstrated by neuropsychological and neuroimaging evidence that indicates distinct neurocognitive networks for semantic knowledge versus tool use and technical reasoning skills (Buxbaum, Schwartz, & Carew, 1997; Goldenberg & Spatt, 2009; Hodges, Bozeat, Lambon Ralph, Patterson, & Spatt, 2000; Osiurak et al., 2009). Evidence also comes from the developmental psychology literature, which has documented that the understanding of physical principles (e.g., support) evolves in infants, although they are obviously unable to generate explanations about them (Baillargeon, 2008; Baillargeon & Hanko-Summers, 1990; Baillargeon, Needham, & Devos, 1992; Needham & Baillargeon, 1993). Thus, infants less than 6 months old infer that any amount of contact between an object and a support is enough for the object to be supported. Between 6 and 9 months, infants improve their understanding and expect an object to be stable if the support is in contact with a significant portion of its surface and begin to understand other features such as the object mass distribution. Explicitly, an object placed on a support is stable if a perpendicular line drawn through the object's center of gravity falls within the support boundaries. Adults adhere to this principle in their predictions, although they

are also generally unaware of it (Needham & Bailargeon, 1993).

In broad terms, the present findings suggest that people can improve a technology over generations without being able to elaborate on the hidden causes by using scientific language. These findings mirror our historical trajectory, through which our ancestors have developed complex technologies over time while living in preliterate societies in which there was no formal pedagogy. Thus, as some have argued, explicit causal understanding is not necessary for CTC (Derex et al., 2019; Harris, Boyd, & Wood, 2021). However, our findings also indicate that explicit causal understanding (meaning using linguistic terms to verbalize what we understand implicitly, i.e., explanatory skills) must not be conflated with causal understanding. Harris et al. (2021) recently provided an anthropological report in which they interviewed skilled Hadza bowyers and compared their bow-making beliefs with those revealed by experimental and engineering research. Hadza bowyers scored better than chance on 7 of 13 questions. Harris et al. (2021) concluded that causal understanding plays a weak role in the evolution of traditional technology. Although they recognized that their interview-based method could have failed to detect “implicit causal knowledge,” they minimized this limitation, emphasizing that this is “possible (though unlikely)” (p. e2). Yet our findings, as already reported by Proffitt et al. (1990), stress that this is likely and that their conclusion should be limited to explicit causal understanding (explanatory skills). The present findings can help reconcile the different approaches by stressing that implicit causal understanding, but not explicit causal understanding, drives CTC.

Our second question (Q2) was whether explanation transmission reinforces the presence of traditions. The similarity analysis failed to support this possibility. Although the within-chain versus between-chain similarity difference was higher in the configurations+theory condition than in other conditions, this difference was not higher, for instance, than the one reported in the configurations condition of the same study (Derex et al., 2019). We also reported that the only condition in which no tradition emerged (no within-chain vs. between-chain similarity difference) was the speed-only condition, which can be explained by the fact that it was the only condition in

which no information about the configurations was transmitted (neither the last two configurations nor explanations). The fact that traditions can emerge without explanation transmission is anything but surprising, given that we do know that cultural phenomena are not restricted to humans (Whiten, 2019, 2021; Whiten, Horner, & De Waal, 2005; Whiten et al., 1999) and, as a result, to the only species able to generate explanations. Nevertheless, it could be expected that explanation transmission reinforces the presence of traditions in chains. Perhaps such a reinforcement is more visible when the technology is more opaque or when all the materials needed to build it are not available, as discussed above. Future research is needed to test this hypothesis. We also did not report any canalizing effect of explanation transmission on exploration, as suggested elsewhere (Derex et al., 2019). The participants did not explore one dimension (e.g., inertia) further than the other because they inherited relevant instructions about that dimension. This canalizing effect was initially proposed (Derex et al., 2019) based on evidence that children who are told the function of a toy tend to explore less and are less likely to discover alternative functions than children who did not receive any information about the toy’s function (Bonawitz et al., 2011; Wood, Kendal, & Flynn, 2015). In these studies, children are influenced to adopt a design stance in Dennett’s terms (Dennett, 1987), because of the focus on the toy’s function. By contrast, our findings stress that participants seemed to take a physical stance when explaining how the wheel system works. Perhaps the canalizing effect of explanation transmission on exploration is greater when people are led to adopt a design stance than a physical stance, a possibility that deserves further investigation.

Before concluding, we would like to delimit the scope of our findings. We focused on written explanations because this form of explanation provides good experimental control. However, written explanations are rigid because they do not allow interactions between the person who generates the explanation and the one who benefits from it, as is the case for oral explanations. This methodological choice necessarily tempers our conclusions, which can nevertheless be more easily applied to everyday contexts in which written instructions can be found (e.g., user instructions). Our findings remain silent as

to several related research questions. For example, an outstanding question is whether people with a good causal understanding of a physical system can make a better use of the explanations they receive about this system than those with a lower understanding of it. Likewise, can improper or proper use of causal explanations hinder or facilitate the development of causal understanding? Exploring these questions requires an experimental design in which participants' causal understanding is assessed both before and after they interact with the physical system to investigate potential evolutions. This experimental design was not used in the present studies reported. Developing such a design might be fruitful to tackle those questions.

To sum up, the key finding is that although the quality of explanations (when analyzed as instructions) increased over generations, we failed to demonstrate that this increase had a direct impact on the performance, bringing into question the role of explanation transmission in CTC. We also noticed that the nature of explanations did not differ according to whether the explanations were generated with the goal of being transmitted (the configurations+theory condition) or not (the speed-only and configurations+noise conditions). These findings remind us that explanations not only are transactional but have also a useful role for the person who generates the explanations, by potentially highlighting an incomplete understanding (Keil, 2006; Keil et al., 2004). This metacognitive function of explanation can take place when the

person is alone but also when they communicate with conspecifics, including in teaching conditions in which the feedback coming from learners can be very useful (for discussion about metacognition and teaching, see Caldwell, Renner, & Atkinson, 2018; Dunstone & Caldwell, 2018; Heyes, 2016; Osiurak & Reynaud, 2020; Shea et al., 2014). In other words, explanatory skills might have a limited impact on the social transmission of technical content but could be particularly beneficial for the individual as a way of increasing their own understanding of a phenomenon (Brockbank, Lombrozo, Gopnik, & Walker, 2023; Gopnik, 1998, 2000; Legare, 2014; Lombrozo, 2016). The question that arises is whether the production of explanations has a greater impact on the model's understanding than on the learner's understanding. Future studies are needed to explore this question, notably in conditions in which participants interact together directly. This counterintuitive perspective also leads us to revisit the payoff generally associated with teaching. Teaching is considered costly for the model, at least in the short term (Caro & Hauser, 1992; Kline, 2015; Kline, Boyd, & Henrich, 2013). However, teaching could also be beneficial in the short term, by potentially increasing the model's understanding. This might be true for explanation-based teaching compared with other forms of teaching (e.g., protolanguage). In sum, the study of explanations can open new avenues in the research on cumulative culture by reminding us that teachers can also learn from learners (Gweon, 2021).

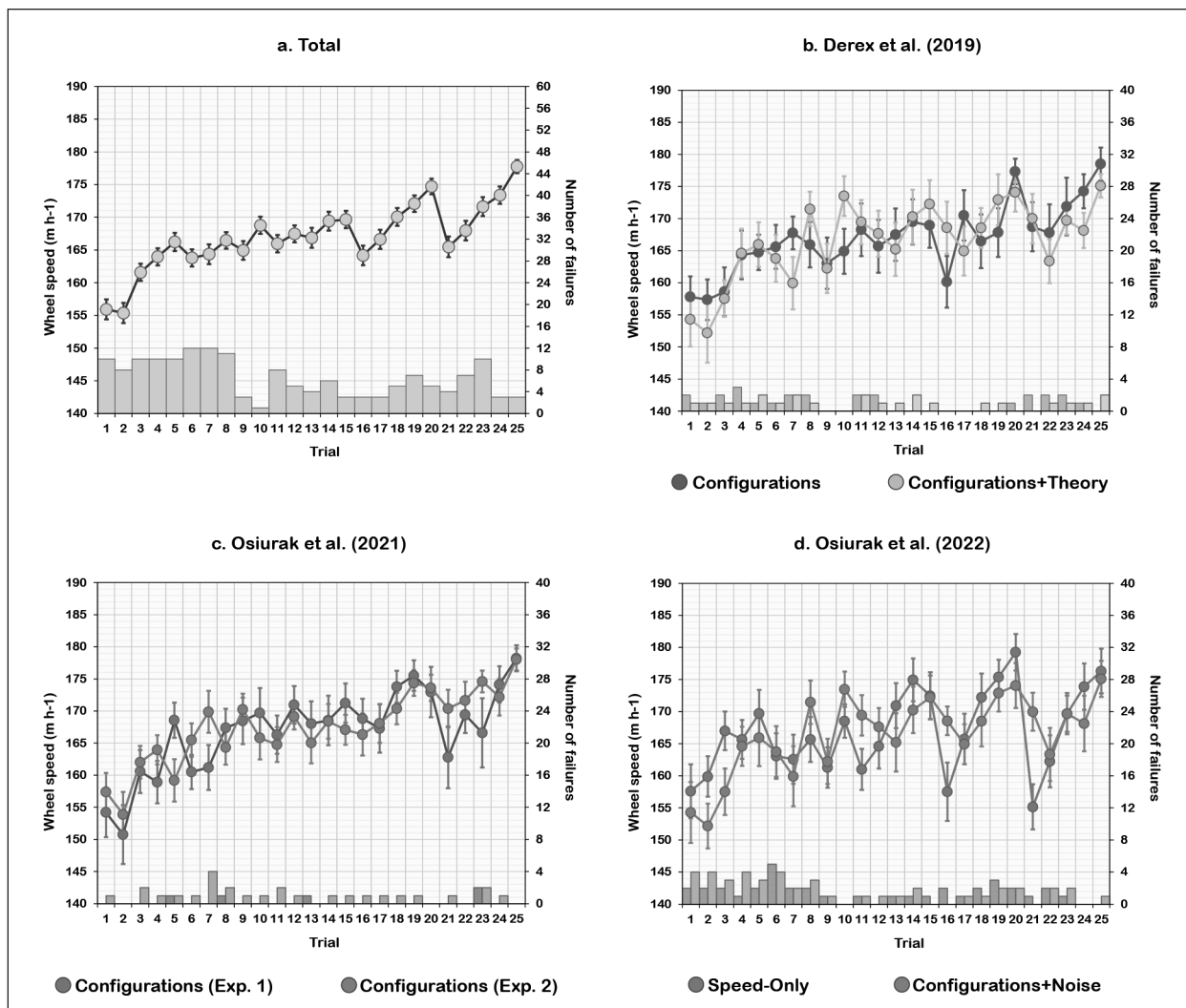


FIGURE S1. Wheel speed over generations for nonfailure wheels and number of failures. Error bars indicate standard errors.

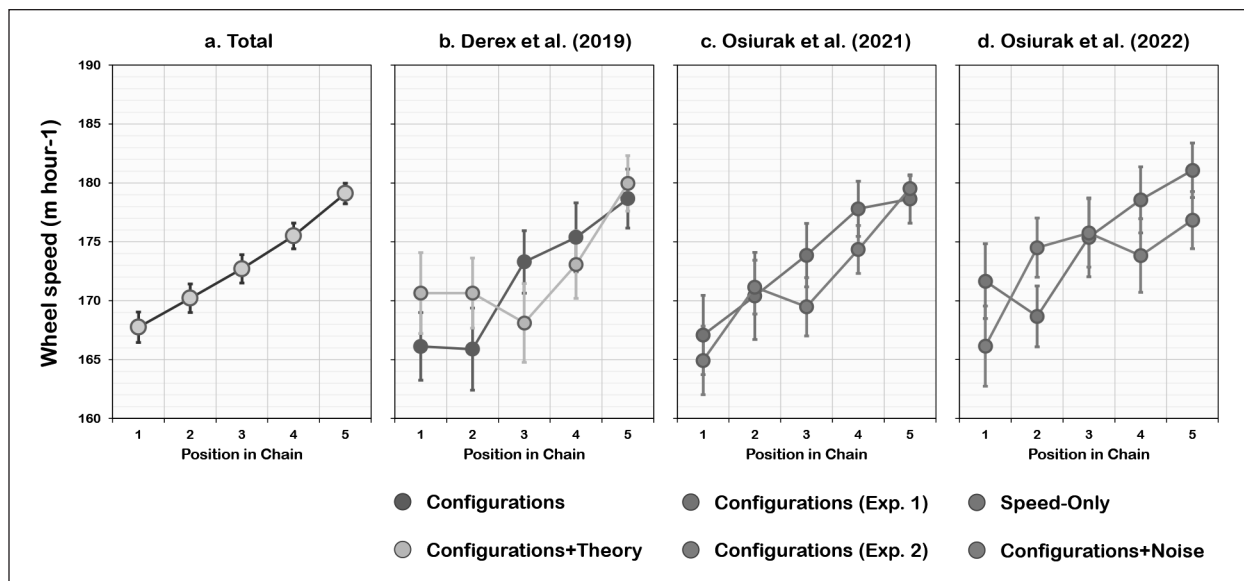


FIGURE S2. Wheel speed (the best speed of the last two trials) over generations. Error bars indicate standard errors

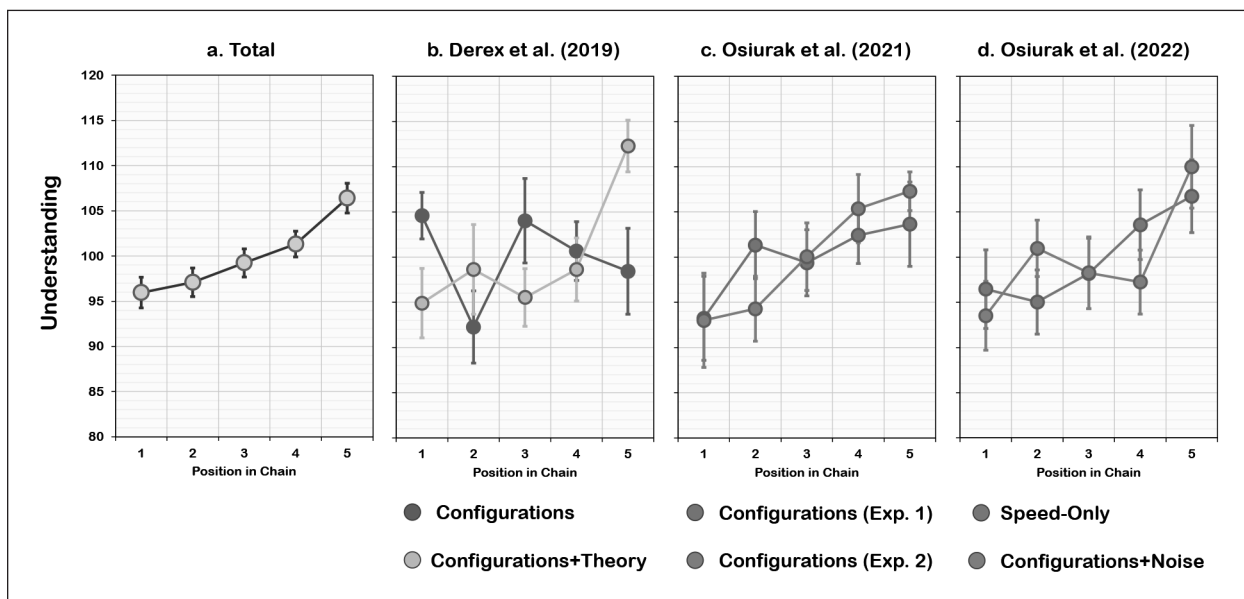


FIGURE S3. Understanding over generations. Error bars indicate standard errors

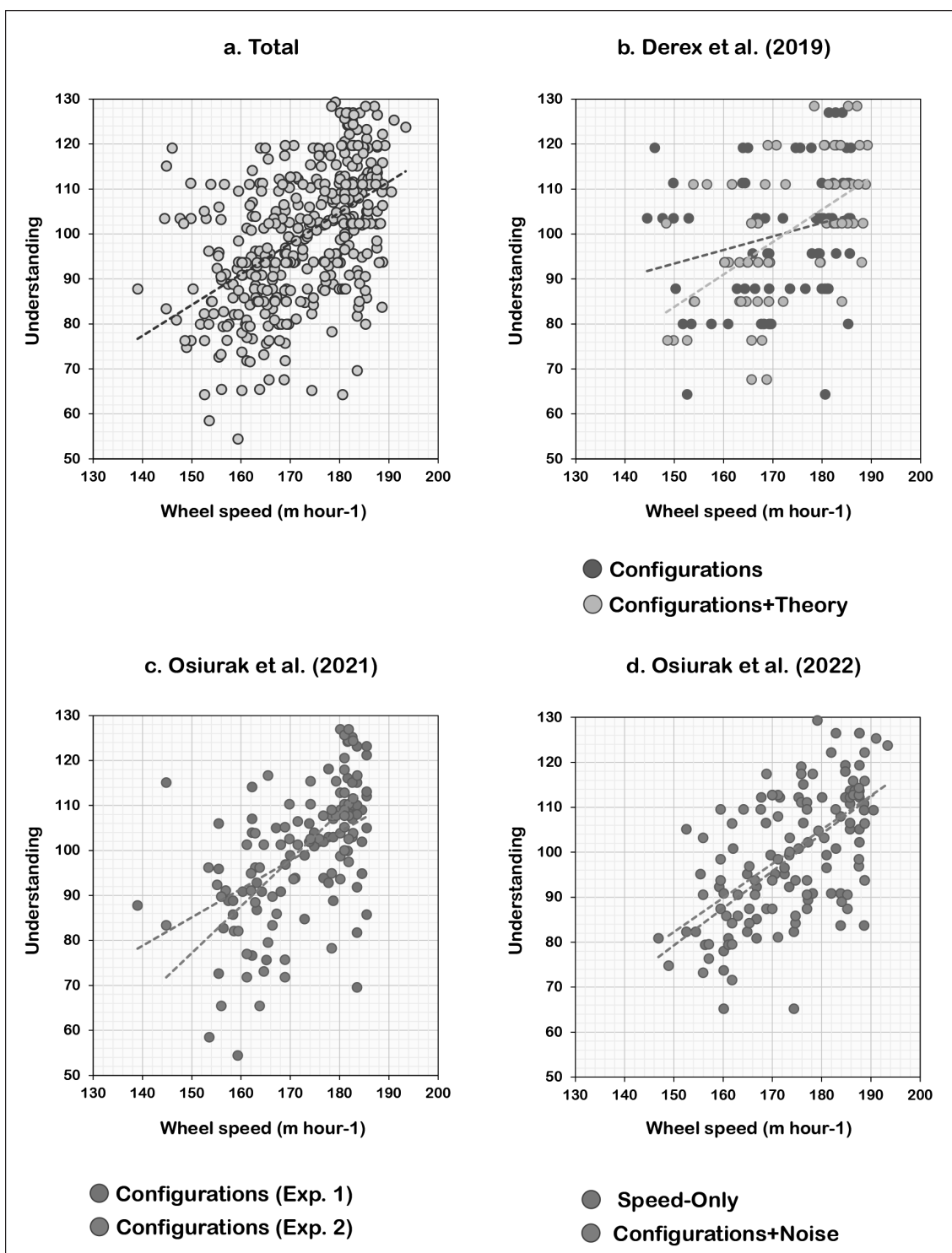


FIGURE S4. Link between wheel speed (the best speed of last two trials) and understanding

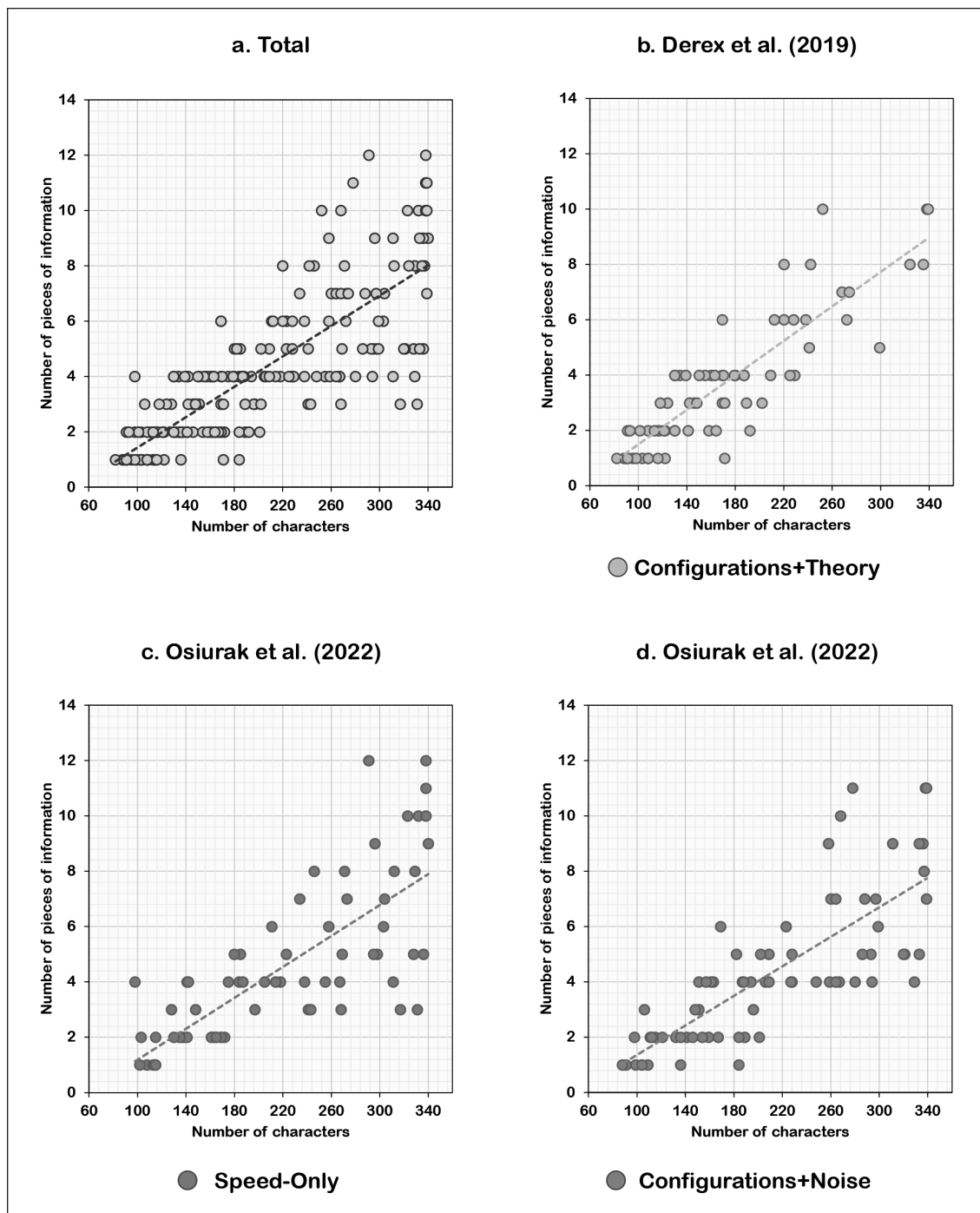


FIGURE S5. Link between number of pieces of information and number of characters

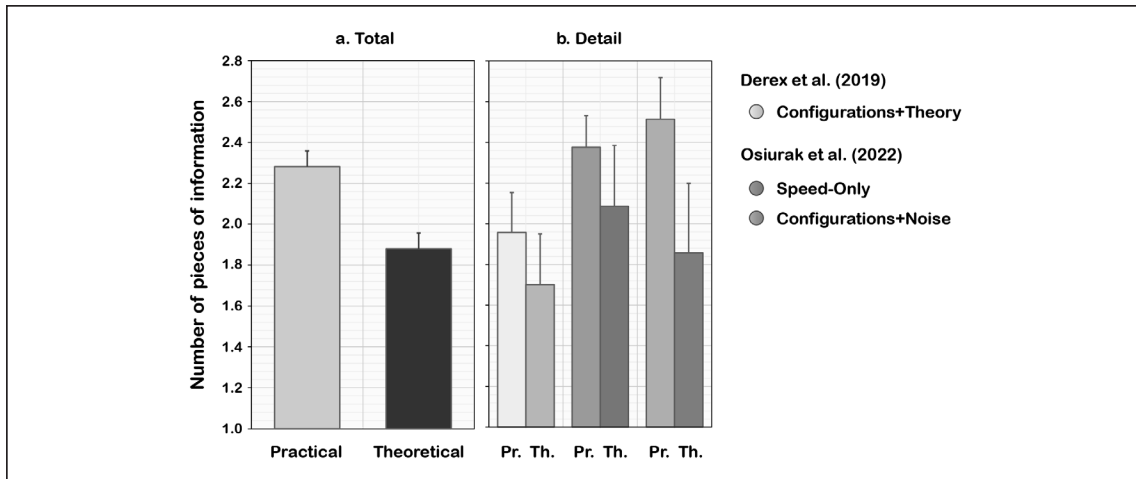


FIGURE S6. Number of pieces of practical (Pr.) versus theoretical (Th.) information. Error bars indicate standard error

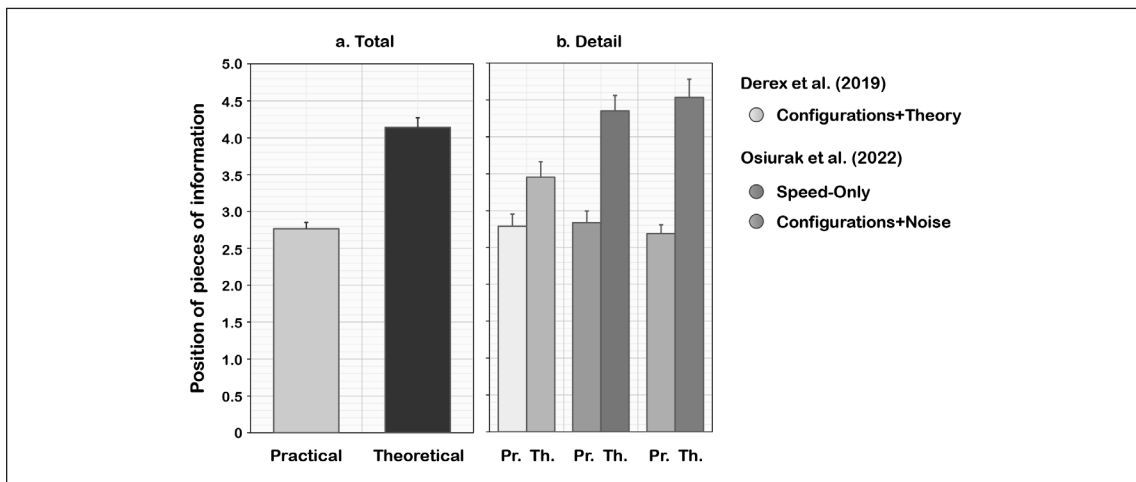


FIGURE S7. Position of pieces of practical (Pr.) versus theoretical (Th.) information. Error bars indicate standard errors

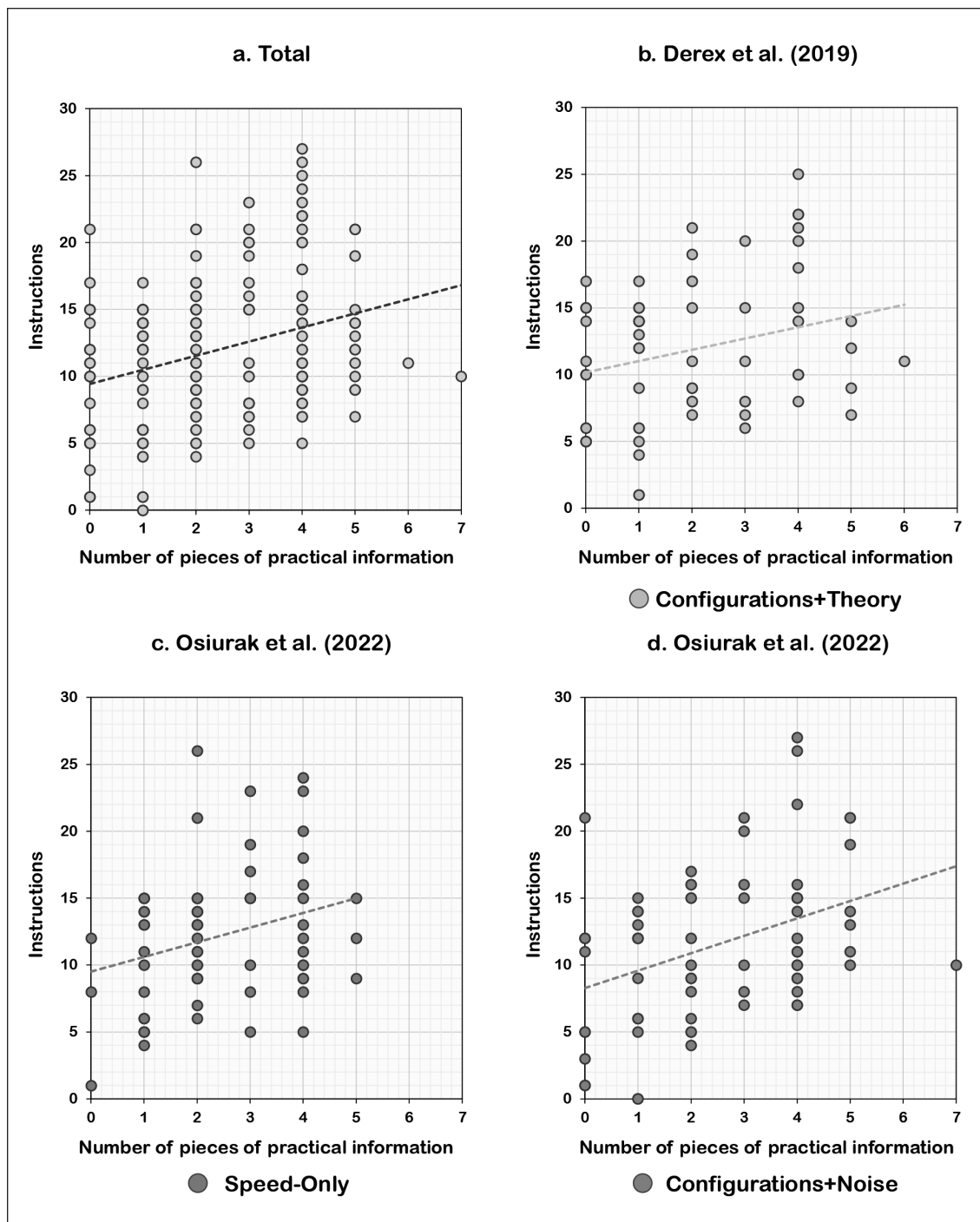


FIGURE S8. Link between instructions and number of pieces of practical information

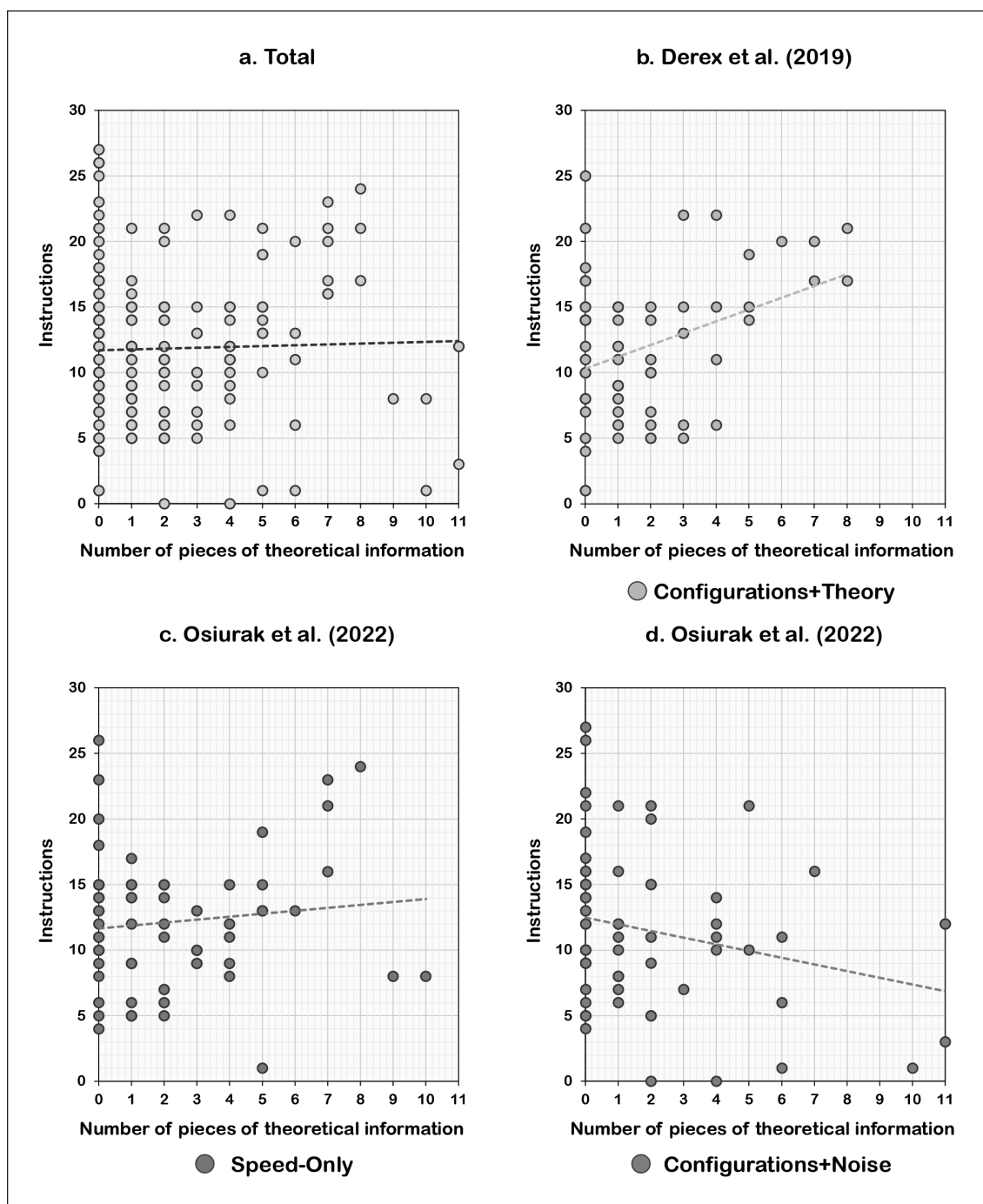


FIGURE S9. Link between instructions and number of pieces of theoretical information

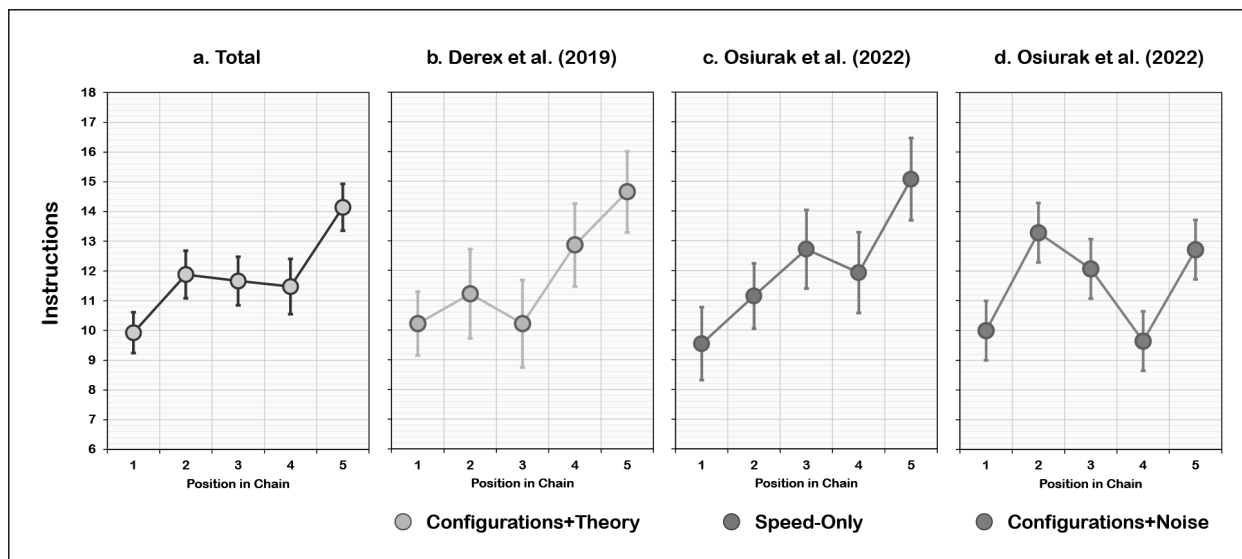


FIGURE S10. Instructions score over generations. Error bars indicate standard errors

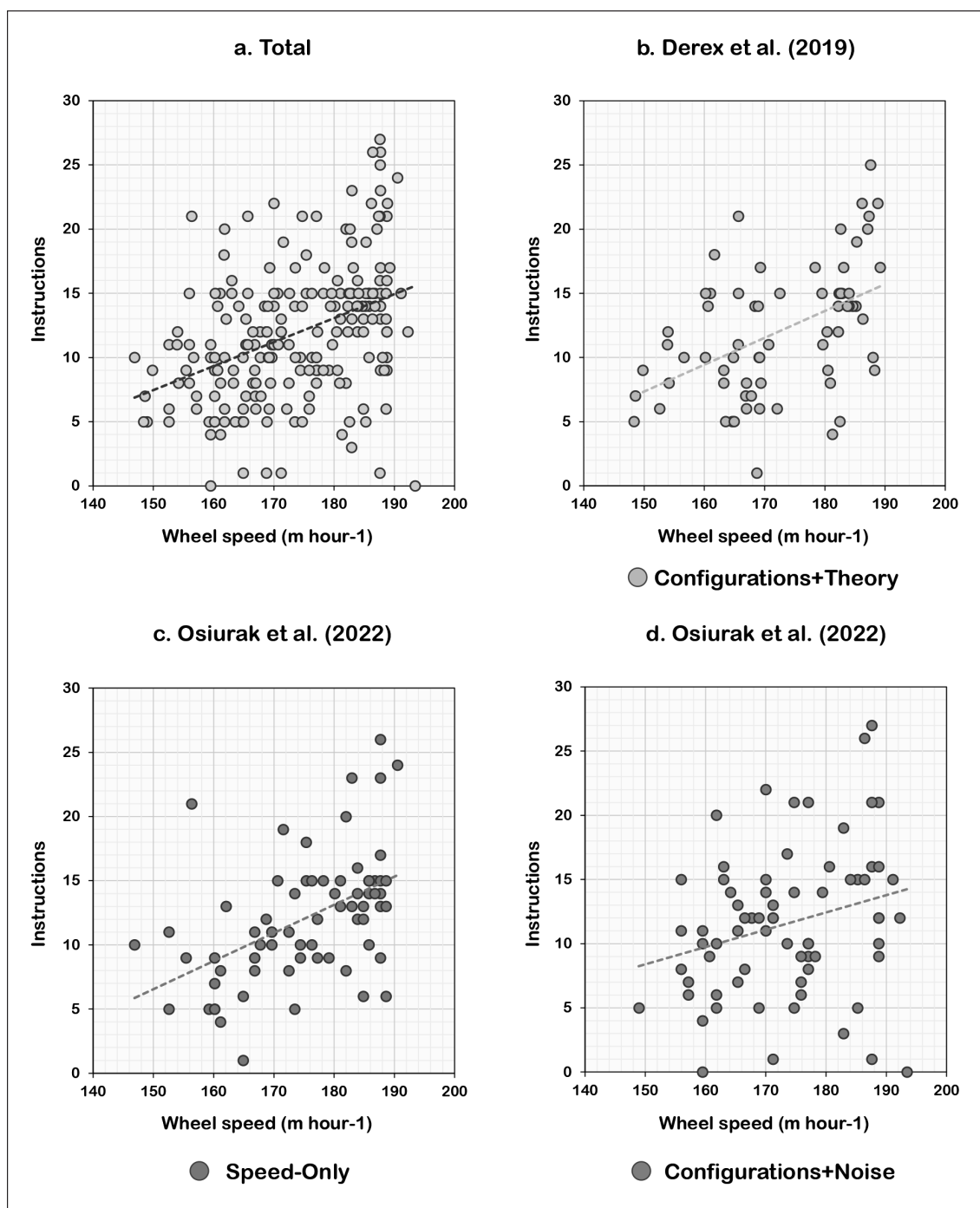


FIGURE S11. Link between instructions and wheel speed (the best of the last two trials)

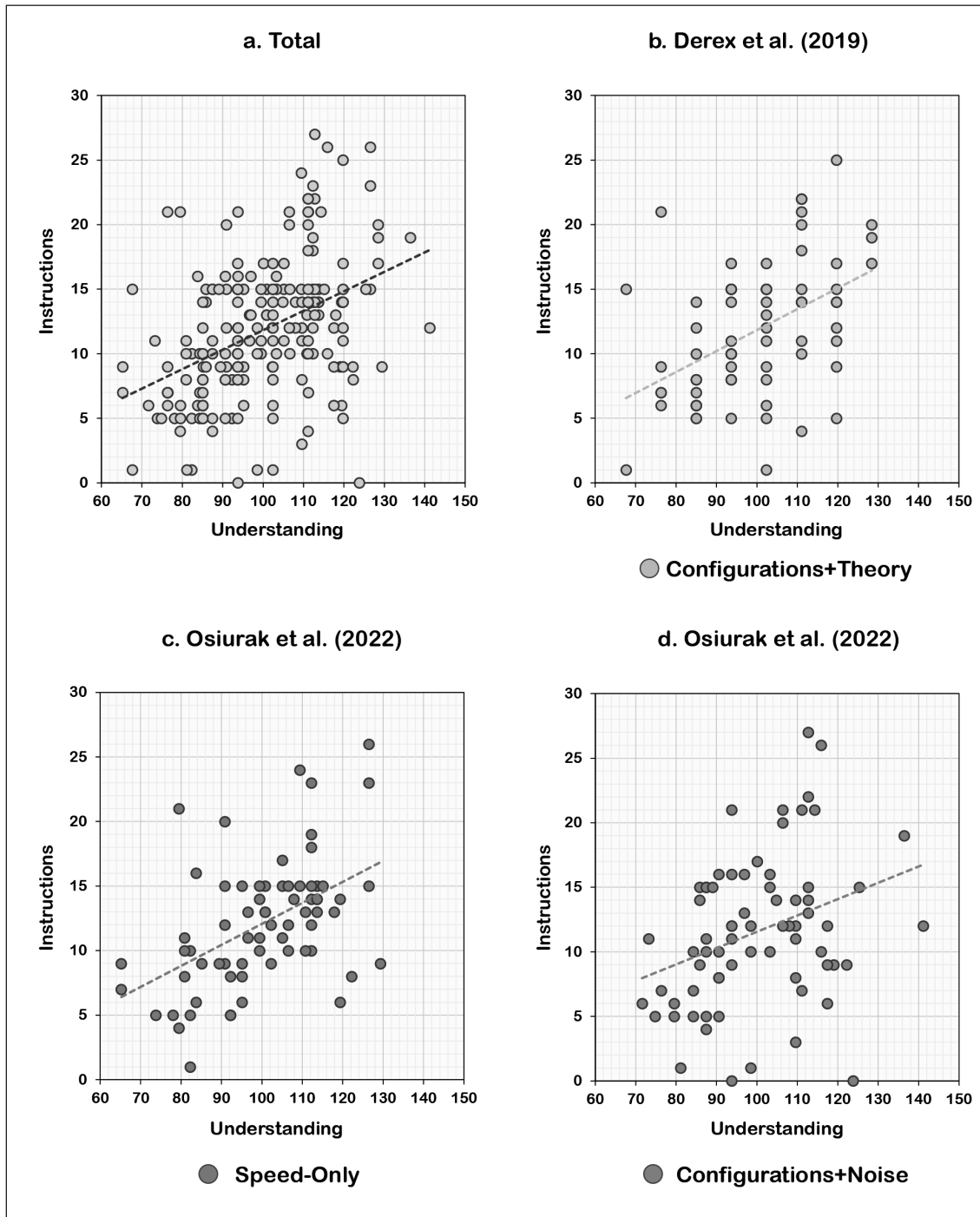


FIGURE S12. Link between instructions and understanding

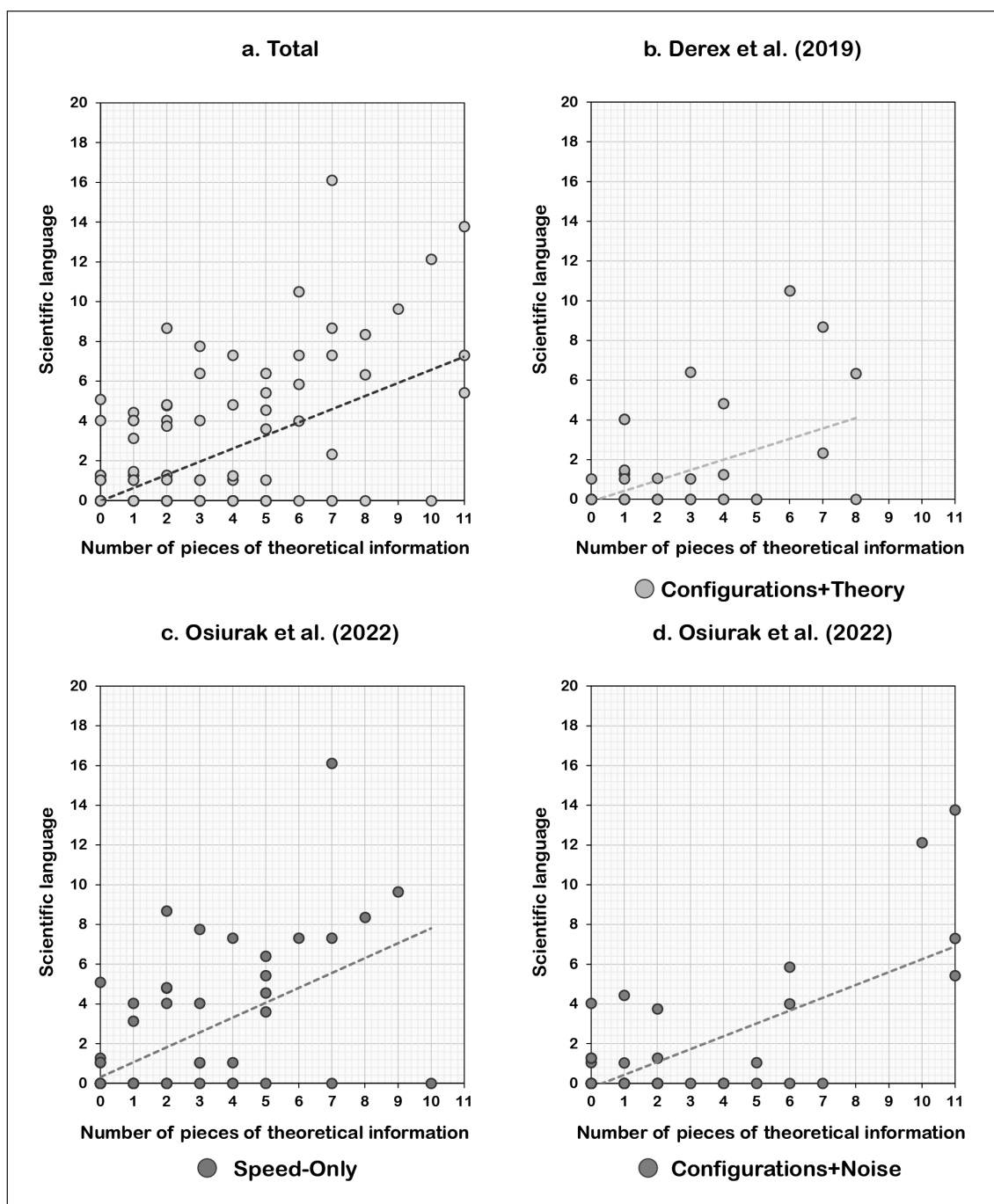


FIGURE S13. Link between scientific language and number of pieces of theoretical information

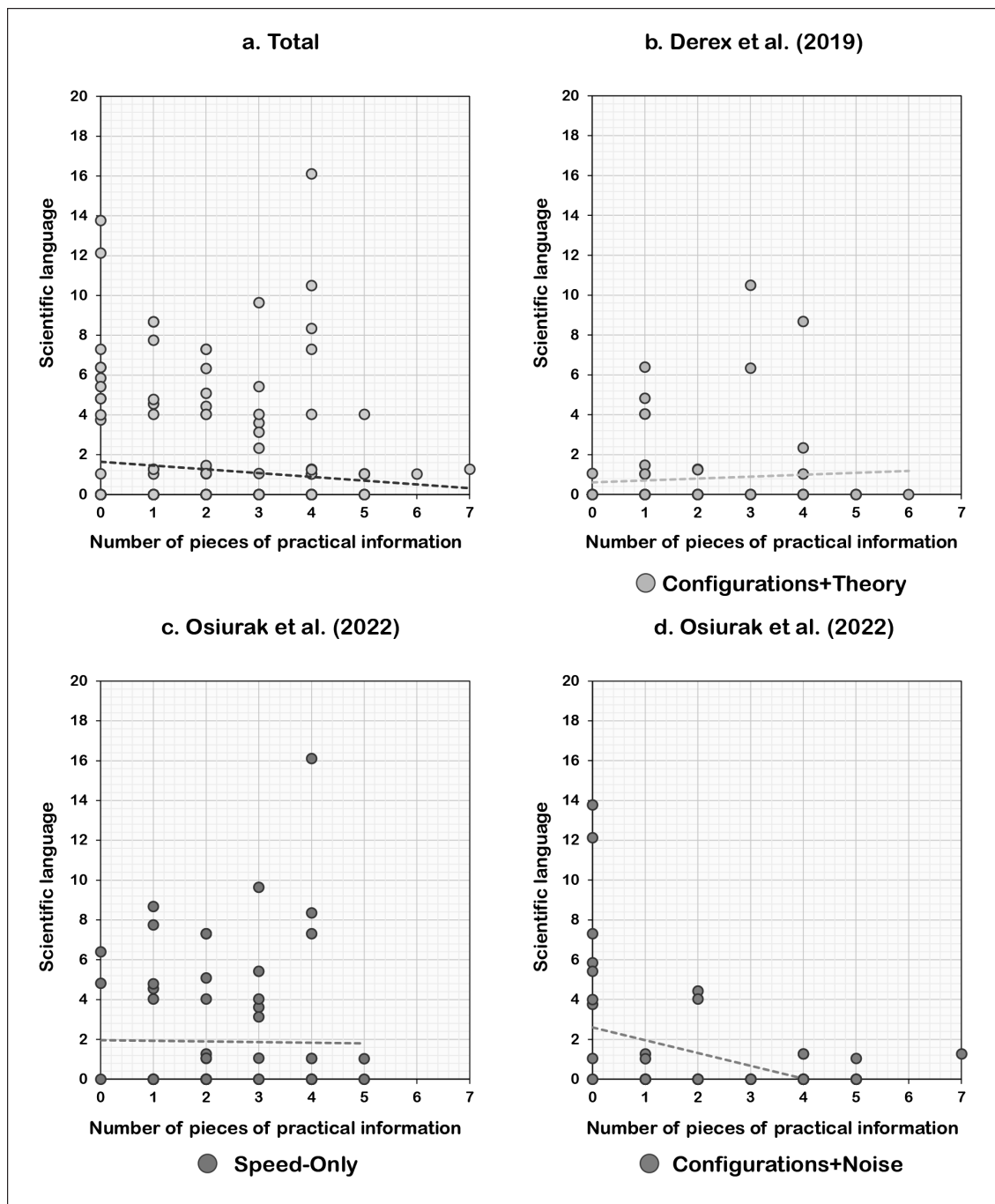


FIGURE S14. Link between scientific language and number of pieces of practical information

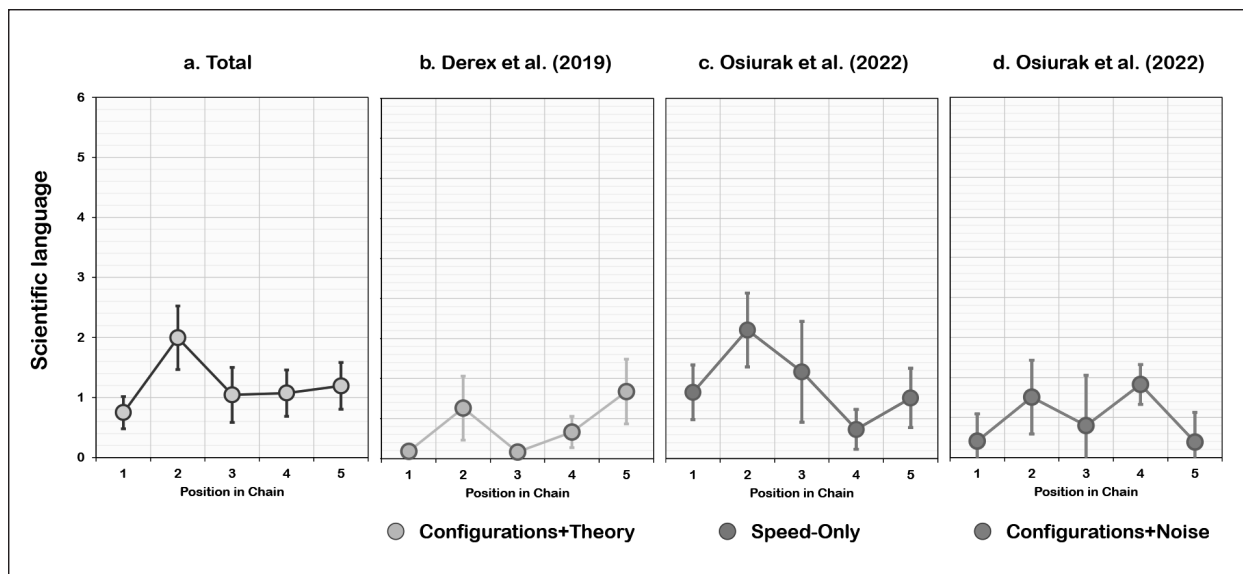


FIGURE S15. Scientific language score over generations. Error bars indicate standard errors

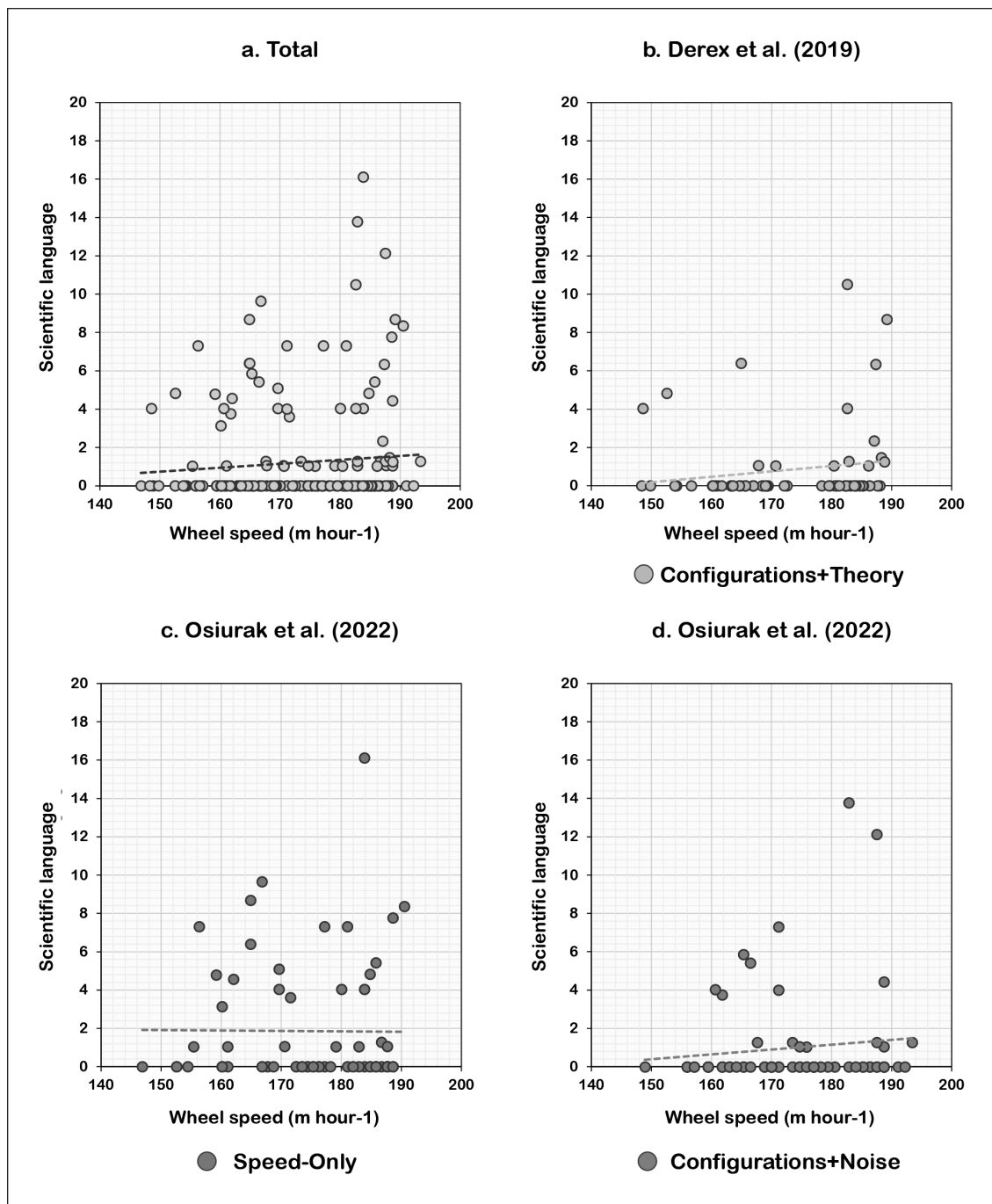


FIGURE S16. Link between scientific language and wheel speed (the best of the last two trials)

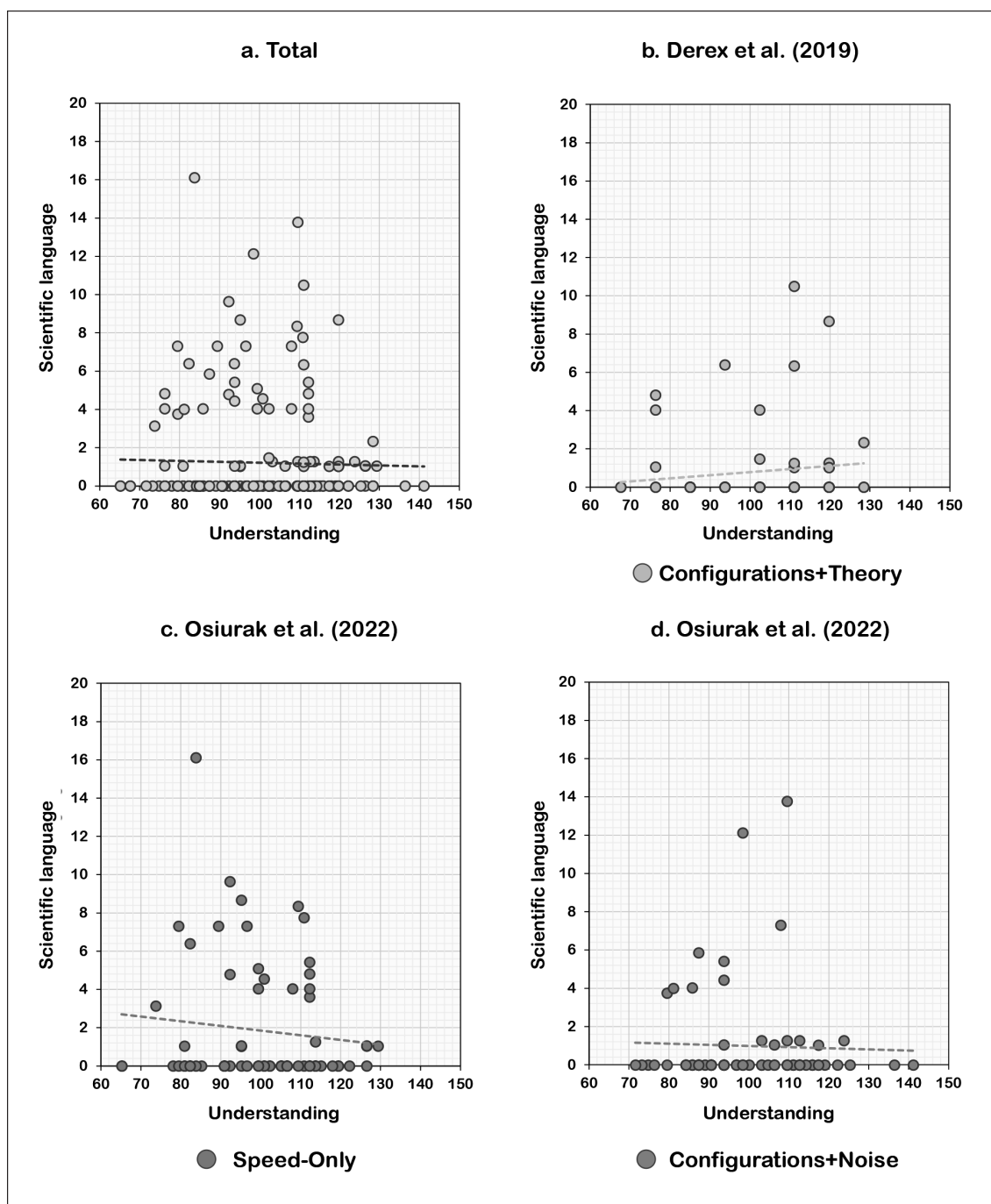


FIGURE S17. Link between scientific language and understanding

TABLE S1. Wheel Speed as a Function of Generation

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	164.58	1.19	161.95	167.02	<.001
Generation	2.82	0.30	2.12	3.39	<.001

Model. Wheel speed ~ Generation + (1|Condition) + (1|Chain's identity).

TABLE S2. Understanding as a Function of Generation

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	92.47	1.65	89.07	95.85	<.001
Generation	2.51	0.46	1.56	3.49	<.001

Model. Understanding ~ Generation + (1|Condition) + (1|Chain's identity).

TABLE S3. Understanding as a Function of Wheel Speed

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	-19.51	10.09	-38.18	2.83	.053
Wheel speed	0.69	0.06	0.56	0.80	<.001

Model. Understanding ~ Wheel speed + (1|Condition) + (1|Generation) + (1|Chain's identity).

TABLE S4. Number of Pieces of Information as a Function of Number of Characters

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	-1.43	0.34	-2.13	-0.77	<.001
Number of characters	0.03	0.01	0.02	0.03	<.001

Model. Number of pieces of information ~ Number of characters + (1|Condition).

TABLE S5. Number of Pieces of Information as a Function of Type of Information

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	2.30	0.15	2.00	2.59	<.001
Theoretical information	-0.40	0.19	-0.78	-0.03	.036

Model. Number of pieces of information ~ Type of information + (1|Condition) + (1|Participant's identity) + (1|Number of characters).

TABLE S6. Position of the Piece of Information as a Function of Type of Information

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	2.61	0.15	2.27	2.93	<.001
Theoretical information	1.08	0.16	0.76	1.39	<.001

Model. Position of the piece of information ~ Type of information + (1|Condition) + (1|Participant's identity).

TABLE S7. Instructions Score as a Function of Number of Pieces of Practical Information

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	9.60	0.86	7.97	11.29	<.001
Practical information	0.98	0.22	0.58	1.40	<.001

Model. Instructions score ~ Number of pieces of practical information + (1|Condition) + (1|Generation) + (1|Chain's identity).

TABLE S8. Instructions Score as a Function of Number of Pieces of Theoretical Information

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	11.80	0.79	10.17	13.44	<.001
Theoretical information	0.01	0.14	-0.27	0.30	.93

Model. Instructions score ~ Number of pieces of theoretical information + (1|Condition) + (1|Generation) + (1|Chain's identity).

TABLE S9. Instructions Score as a Function of Generation

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	9.45	0.85	7.86	10.90	<.001
Generation	0.79	0.24	0.30	1.30	<.001

Model. Instructions score ~ Generation + (1|Condition) + (1|Chain's identity).

TABLE S10. Instructions Score as a Function of Wheel Speed

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	-21.20	5.38	-32.19	-10.26	<.001
Wheel speed	0.19	0.03	0.13	0.25	<.001

Model. Instructions score ~ Wheel speed + (1|Generation) + (1|Condition) + (1|Chain's identity).

TABLE S11. Instructions Score as a Function of Understanding

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	−3.05	2.31	−7.52	1.41	.19
Understanding	0.15	0.02	0.10	0.19	<.001

Model. Instructions score ~ Understanding + (1|Generation) + (1|Condition) + (1|Chain's identity).

TABLE S12. Scientific Language Score as a Function of Number of Pieces of Theoretical Information

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	−0.01	0.31	−0.69	0.67	.97
Theoretical information	0.65	0.06	0.54	0.77	<.001

Model. Scientific language score ~ Number of pieces of theoretical information + (1|Condition) + (1|Generation) + (1|Chain's identity).

TABLE S13. Scientific Language Score as a Function of Number of Pieces of Practical Information

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	1.73	0.46	0.79	2.69	.012
Practical information	−0.23	0.12	−0.45	0.01	.057

Model. Scientific language score ~ Number of pieces of practical information + (1|Condition) + (1|Generation) + (1|Chain's identity).

TABLE S14. Scientific Language Score as a Function of Generation

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	1.22	0.50	0.25	2.19	.034
Generation	−0.01	0.13	−0.25	0.25	.98

Model. Scientific language score ~ Generation + (1|Condition) + (1|Chain's identity).

TABLE S15. Scientific Language Score as a Function of Wheel Speed

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	−1.91	2.94	−7.77	3.79	.52
Wheel speed	0.02	0.02	−0.01	0.05	.29

Model. Scientific language score ~ Wheel speed + (1|Generation) + (1|Condition) + (1|Chain's identity).

TABLE S16. Scientific Language Score as a Function of Understanding

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	1.84	1.31	−0.73	4.38	.16
Understanding	−0.01	0.01	−0.03	0.02	.62

Model. Scientific language score ~ Understanding + (1|Generation) + (1|Condition) + (1|Chain's identity).

TABLE S17. Similarity Score as a Function of Type of Similarity and Condition

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	28.34	0.89	26.54	29.90	<.001
TypeWithin	5.51	0.62	4.35	6.71	<.001
ConditionConfigurations+Theory	0.81	1.15	−1.62	2.94	.48
ConditionConfigurationsExp1	2.40	1.15	0.27	4.87	.04
ConditionConfigurationsExp2	−0.52	1.15	−2.52	1.76	.65
ConditionConfigurations+Noise	−1.85	1.15	−0.34	4.19	.11
ConditionSpeed-Only	−1.08	1.15	−3.28	1.09	.35
TypeWithin:ConditionConfigurations+Theory	−0.40	0.88	−2.16	1.24	.65
TypeWithin:ConditionConfigurationsExp1	−2.63	0.88	−4.24	−1.05	.003
TypeWithin:ConditionConfigurationsExp2	−1.75	0.88	−3.47	−0.28	.047
TypeWithin:ConditionConfigurations+Noise	−3.03	0.88	−4.84	−1.32	<.001
TypeWithin:ConditionSpeed-Only	−5.87	0.88	−7.61	−4.18	<.001

Model. Similarity score ~ Type of similarity*Condition + (1|Generation) + (1|Chain's identity).

TABLE S18. Inertia Exploration Score as a function of Inertia Instructions Score Inherited

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	14.88	2.59	9.94	19.81	<.001
Inertia_Inherited	−0.21	0.33	−0.85	0.42	.51
ConditionConfigurations+Theory	−3.61	3.22	−9.90	2.62	.26
ConditionSpeed-Only	12.42	3.55	5.40	19.22	<.001
Inertia_Inherited:ConditionConfigurations+Theory	0.61	0.45	−0.26	1.49	.18
Inertia_Inherited:ConditionSpeed-Only	0.18	0.45	−0.68	1.08	.69

Model. Inertia exploration score ~ Inertia instructions score inherited*Condition + (1|Generation) + (1|Chain's identity).

TABLE S19. Inertia Exploration Score as a Function of Center-of-Mass Instructions Score Inherited

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	17.45	2.79	12.15	22.78	<.001
CoM_Inherited	−0.61	0.33	−1.27	0.02	.07
ConditionConfigurations+Theory	−3.54	3.47	−10.34	3.11	.31
ConditionSpeed-Only	7.39	3.34	0.89	13.85	<.001
CoM_Inherited:ConditionConfigurations+Theory	0.53	0.47	−0.37	1.44	.26
CoM_Inherited:ConditionSpeed-Only	1.14	0.50	0.16	2.12	.026

Model. Inertia exploration score ~ Center-of-mass instructions score inherited*Condition + (1|Generation) + (1|Chain's identity).

TABLE S20. Center-of-Mass Exploration Score as a Function of Inertia Instructions Score Inherited

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	8.90	2.43	4.31	13.50	<.001
Inertia_Inherited	−0.01	0.32	−0.62	0.61	.99
ConditionConfigurations+Theory	−3.25	3.44	−9.86	3.29	.35
ConditionSpeed-Only	−0.11	3.64	−7.03	7.03	.97
Inertia_Inherited:ConditionConfigurations+Theory	−0.32	0.46	−1.19	0.56	.48
Inertia_Inherited:ConditionSpeed-Only	−0.09	0.44	−0.97	0.74	.83

Model. Center-of-mass exploration score ~ Inertia instructions score inherited*Condition + (1|Generation) + (1|Chain's identity).

TABLE S21. Center-of-Mass Exploration Score as a Function of Center-of-Mass Instructions Score Inherited

	Estimate	SD	L95%	U95%	<i>p</i>
Intercept	9.42	2.68	4.22	14.61	<.001
CoM_Inherited	−0.09	0.33	−0.74	0.57	.79
ConditionConfigurations+Theory	−6.24	3.68	−13.31	1.08	.09
ConditionSpeed-Only	−1.47	3.52	−8.39	5.34	.68
CoM_Inherited:ConditionConfigurations+Theory	0.21	0.47	−0.74	1.14	.65
CoM_Inherited:ConditionSpeed-Only	0.12	0.50	−0.83	1.17	.80

Model. Center-of-mass exploration score ~ Center-of-mass instructions score inherited*Condition + (1|Generation) + (1|Chain's identity).

NOTES

This work was supported by grants from Région Auvergne-Rhône-Alpes (NUMERICOG-2017-900-EA 3082 EMC-R-2075; F.O.) and the French National Research Agency (ANR; Project TECHNITION: ANR-21-CE28-0023-01; F.O. and N.C.).

The data that support the findings of the study and the codes used in this article are available at <https://osf.io/yghxu/>.

The authors declare that they have no competing interests.

Address correspondence about this article to François Osiurak, Laboratoire EMC, Université Lyon 2, 5 Avenue Pierre Mendès France, 69676 Bron Cedex, France (email: francois.osiurak@univ-lyon2.fr).

REFERENCES

- Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science*, 3, 2–13. <https://doi.org/10.1111/j.1745-6916.2008.00056.x>
- Baillargeon, R., & Hanko-Summers, S. (1990). Is the top object adequately supported by the bottom object? Young infants' understanding of support relations. *Cognitive Development*, 5, 29–53. [https://doi.org/10.1016/0885-2014\(90\)90011-H](https://doi.org/10.1016/0885-2014(90)90011-H)
- Baillargeon, R., Needham, A., & Devos, J. (1992). The development of young infants' intuitions about support. *Early Development and Parenting*, 1, 69–78. <https://doi.org/10.1002/edp.2430010203>
- Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, 120, 322–330. <https://doi.org/10.1016/j.cognition.2010.10.001>
- Boyd, R., & Richerson, P. J. (1995). Why culture is common, but cultural evolution is rare. *Proceedings of the British Academy*, 88, 77–93.
- Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 10918–10925. <https://doi.org/10.1073/pnas.1100290108>
- Brockbank, E., Lombrozo, T., Gopnik, A., & Walker, C. M. (2023). Ask me why, don't tell me why : Asking children for explanations facilitates relational thinking. *Developmental Science*, 26, e13274. <https://doi.org/10.1111/desc.13274>
- Buxbaum, L. J., Schwartz, M. F., & Carew, T. G. (1997). The role of semantic memory in object use. *Cognitive Neuropsychology*, 14, 219–254. <https://doi.org/10.1080/026432997381565>
- Caldwell, C. A., Renner, E., & Atkinson, M. (2018). Human teaching and cumulative cultural evolution. *Review of Philosophy and Psychology*, 9, 751–770. <https://doi.org/10.1007/s13164-017-0346-3>
- Caro, T. M., & Hauser, M. D. (1992). Is there teaching in nonhuman animals? *Quarterly Review of Biology*, 67, 151–174. <https://doi.org/10.1086/417553>
- Dennett, D. C. (1987). *The intentional stance*. MIT Press.
- Dere, M. (2022). Human cumulative culture and the exploitation of natural phenomena. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1843), 20200311. <https://doi.org/10.1098/rstb.2020.0311>
- Dere, M., Bonnefon, J. F., Boyd, R., & Mesoudi, A. (2019). Causal understanding is not necessary for the improvement of culturally evolving technology. *Nature Human Behaviour*, 3(5), 446–452. <https://doi.org/10.1038/s41562-019-0567-9>
- Dunstone, J., & Caldwell, C. A. (2018). Cumulative culture and explicit metacognition: A review of theories, evidence and key predictions. *Palgrave Communications*, 4(1), 1–11. <https://doi.org/10.1057/s41599-018-0200-y>
- Gardenfors, P., & Höglberg, A. (2017). The archaeology of teaching and the evolution of *Homo docens*. *Current Anthropology*, 58(2), 188–208. <https://doi.org/10.1086/691178>
- Goldenberg, G., & Spatt, J. (2009). The neural basis of tool use. *Brain*, 132, 1645–1655. <https://doi.org/10.1093/brain/awp080>
- Gopnik, A. (1998). Explanation as orgasm. *Minds and Machines*, 8(1), 101–118. <https://doi.org/10.1023/A:1008290415597>
- Gopnik, A. (2000). Explanation as orgasm and the drive for causal understanding: The evolution, function and phenomenology of the theory-formation system. In F. C. Keil & R. Wilson (Eds.), *Cognition and explanation* (pp. 299–323). MIT Press.
- Gopnik, A., & Schulz, L. (2004). Mechanisms of theory formation in young children. *Trends in Cognitive Sciences*, 8, 371–377. <https://doi.org/10.1016/j.tics.2004.06.005>
- Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, 111, 3–32. <https://doi.org/10.1037/0033-295X.111.1.3>
- Gweon, H. (2021). Inferential social learning: Cognitive foundations of human social learning and teaching. *Trends in Cognitive Sciences*, 25(10), 896–910. <https://doi.org/10.1016/j.tics.2021.07.008>
- Harris, J. A., Boyd, R., & Wood, B. M. (2021). The role of causal knowledge in the evolution of traditional technology. *Current Biology*, 31, 1–6. <https://doi.org/10.1016/j.cub.2021.01.096>
- Heyes, C. (1994). Social learning in animals: Categories and mechanisms. *Biological Reviews*, 69, 207–231. <https://doi.org/10.1111/j.1469-185X.1994.tb01506.x>

- Heyes, C. (2016). Who knows? Metacognitive social learning strategies. *Trends in Cognitive Sciences*, 20, 204–213. <https://doi.org/10.1016/j.tics.2015.12.007>
- Heyes, C. (2021a). Imitation. *Current Biology*, 31, R228–R232. <https://doi.org/10.1016/j.cub.2020.11.071>
- Heyes, C. (2021b). Imitation and culture: What gives? *Mind & Language*, 38, 42–63. <https://doi.org/10.1111/mila.12388>
- Hodges, J. R., Bozeat, S., Lambon Ralph, M. A., Patterson, K., & Spatt, J. (2000). The role of conceptual knowledge in object use: Evidence from semantic dementia. *Brain*, 123, 1913–1925. <https://doi.org/10.1093/brain/123.9.1913>
- Jagiello, R., Heyes, C., & Whitehouse, H. (2022). Tradition and invention: The bifocal stance theory of cultural evolution. *Behavioral and Brain Sciences*, 45, e249. <https://doi.org/10.1017/S0140525X22000383>
- Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology*, 57, 227–254. <https://doi.org/10.1146/annurev.psych.57.102904.190100>
- Keil, F. C., Rozenblit, L., & Mills, C. M. (2004). What lies beneath? Understanding the limits of understanding. In D. T. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (pp. 227–249). MIT Press.
- Kline, M. A. (2015). How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behavioral and Brain Sciences*, 38, 1–71. <https://doi.org/10.1017/S0140525X14000090>
- Kline, M. A., Boyd, R., & Henrich, J. (2013). Teaching and the life history of cultural transmission in Fijian villages. *Human Nature*, 24(4), 351–374. <https://doi.org/10.1007/s12110-013-9180-1>
- Kubricht, J. R., Holyoak, K. J., & Lu, H. (2017). Intuitive physics: Current research and controversies. *Trends in Cognitive Sciences*, 21, 749–759. <https://doi.org/10.1016/j.tics.2017.06.002>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Legare, C. H. (2014). The contributions of explanation and exploration to children's scientific reasoning. *Child Development Perspectives*, 8, 101–106. <https://doi.org/10.1111/cdep.12070>
- Lombrozo, T. (2016). Explanatory preferences shape learning and inference. *Trends in Cognitive Sciences*, 20, 748–759. <http://dx.doi.org/10.1016/j.tics.2016.08.001>
- McCloskey, M. (1983). Intuitive physics. *Scientific American*, 248(4), 122–131.
- McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naïve beliefs about the motion of objects. *Science*, 210(4474), 1139–1141. <https://doi.org/10.1126/science.210.4474.1139>
- McCloskey, M., Washburn, A., & Felch, L. (1983). Intuitive physics: The straight-down belief and its origin. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 636–649. <https://doi.org/10.1037/0278-7393.9.4.636>
- Mesoudi, A., & Thornton, A. (2018). What is cumulative cultural evolution? *Proceedings of the Royal Society B: Biological Sciences*, 285(1880), 20180712. <https://doi.org/10.1098/rspb.2018.0712>
- Morgan, T. J. H., Uomini, N. T., Rendell, L. E., Chouinard-Thuly, L., Street, S. E., Lewis, H. M., . . . Laland, K. N. (2015). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nature Communications*, 6, 6029. <https://doi.org/10.1038/ncomms7029>
- Needham, A., & Baillargeon, R. (1993). Intuitions about support in 4.5-month-old infants. *Cognition*, 47, 121–148. [https://doi.org/10.1016/0010-0277\(93\)90002-D](https://doi.org/10.1016/0010-0277(93)90002-D)
- Osiurak, F., & Reynaud, E. (2020). The elephant in the room: What matters cognitively in cumulative technological culture. *Behavioral and Brain Sciences*, 43, e156. <https://doi.org/10.1017/S0140525X19003236>
- Osiurak, F., Claidière, N., & Federico, G. (2023). Bringing cumulative technological culture beyond copying versus reasoning. *Trends in Cognitive Sciences*, 27, 30–42. <https://doi.org/10.1016/j.tics.2022.09.024>
- Osiurak, F., Claidière, N., Bluet, A., Brogniart, J., Lasserre, S., Bonhoure, T., . . . Reynaud, E. (2022). Technical reasoning bolsters cumulative technological culture through convergent transformations. *Science Advances*, 8(9), eabl7446. <https://doi.org/10.1126/sciadv.abl7446>
- Osiurak, F., Cretel, C., Duhau-Marmon, N., Fournier, I., Marignier, L., De Oliveira, E., . . . Reynaud, E. (2020). The pedagogue, the engineer, and the friend: From whom do we learn? *Human Nature*, 31(4), 462–482. <https://doi.org/10.1007/s12110-020-09379-0>
- Osiurak, F., De Oliveira, E., Navarro, J., & Reynaud, E. (2020). The castaway island: Distinct roles of theory of mind and technical reasoning in cumulative technological culture. *Journal of Experimental Psychology: General*, 149, 58–66. <https://doi.org/10.1037/xge0000614>
- Osiurak, F., Jarry, C., Allain, P., Aubin, G., Etcharry-Bouyx, F., Richard, I., . . . Le Gall, D. (2009). Unusual use of objects after unilateral brain damage: The technical reasoning model. *Cortex*, 45, 769–783. <https://doi.org/10.1016/j.cortex.2008.06.013>
- Osiurak, F., Lasserre, S., Arbanti, J., Brogniart, J., Bluet, A., Navarro, J., & Reynaud, E. (2021). Technical reasoning is important for cumulative technological culture. *Nature Human Behaviour*, 5, 1643–1651. <https://doi.org/10.1038/s41562-021-01159-9>
- Pearl, J. (2000). *Causality*. Cambridge University Press.
- Perry, S., Carter, A., Smolla, M., Akçay, E., Nöbel, S., Foster, J. G., & Healy, S. D. (2021). Not by transmission alone:

- The role of invention in cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1828), 20200049. <https://doi.org/10.1098/rstb.2020.0049>
- Proffitt, D. R., Kaiser, M. K., & Whelan, S. M. (1990). Understanding wheel dynamics. *Cognitive Psychology*, 22, 342–373. [https://doi.org/10.1016/0010-0285\(90\)90007-Q](https://doi.org/10.1016/0010-0285(90)90007-Q)
- R Development Core Team. (2011). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Ratinaud, P. (2009). *IRAMUTEQ: Interface de R pour les analyses multidimensionnelles de textes et de questionnaires*. <http://www.iramuteq.org/>
- Reindl, E., & Tennie, C. (2018). Young children fail to generate an additive ratchet effect in an open-ended construction task. *PLoS One*, 13(6), 1–22. <https://doi.org/10.1371/journal.pone.0197828>
- Shea, N., Boldt, A., Bang, D., Yeung, N., Heyes, C., & Frith, C. D. (2014). Supra-personal cognitive control and metacognition. *Trends in Cognitive Sciences*, 18, 186–193. <https://doi.org/10.1016/j.tics.2014.01.006>
- Singh, M., Acerbi, A., Caldwell, C. A., Danchin, É., Isabel, G., Molleman, L., . . . Derex, M. (2021). Beyond social learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1828), 20200050. <https://doi.org/10.1098/rstb.2020.0050>
- Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: On the evolution of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1528), 2405–2415. <https://doi.org/10.1098/rstb.2009.0052>
- Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, 16, 495–511. <https://doi.org/10.1017/S0140525X0003123X>
- Vale, G. L., McGuigan, N., Burdett, E., Lambeth, S. P., Lucas, A., Rawlings, B., . . . Whiten, A. (2021). Why do chimpanzees have diverse behavioral repertoires yet lack more complex cultures? Invention and social information use in a cumulative task. *Evolution and Human Behavior*, 42, 247–258. <https://doi.org/10.1016/j.evolhumbehav.2020.11.003>
- Waldmann, M. R., & Hagmayer, Y. (2005). Seeing versus doing: Two modes of accessing causal knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 216–227. <https://doi.org/10.1037/0278-7393.31.2.216>
- Whiten, A. (2019). Cultural evolution in animals. *Annual Review of Ecology, Evolution, and Systematics*, 50(1), 27–48. <https://doi.org/10.1146/annurev-ecolsys-110218-025040>
- Whiten, A. (2021). The burgeoning reach of animal culture. *Science*, 372(6537), eabe6514. <https://doi.org/10.1126/science.abe6514>
- Whiten, A. (2022). Blind alleys and fruitful pathways in the comparative study of cultural cognition. *Physics of Life Reviews*, 43, 211–238. <https://doi.org/10.1016/j.plrev.2022.10.003>
- Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., . . . Boesch, C. (1999). Cultures in chimpanzees. *Nature*, 399(6737), 682–685. <https://doi.org/10.1038/21415>
- Whiten, A., Horner, V., & De Waal, F. B. M. (2005). Conformity to cultural norms of tool use in chimpanzees. *Nature*, 437(7059), 737–740. <https://doi.org/10.1038/nature04047>
- Wood, L. A., Kendal, R. L., & Flynn, E. G. (2015). Does a peer model's task proficiency influence children's solution choice and innovation? *Journal of Experimental Child Psychology*, 139, 190–202. <https://doi.org/10.1016/j.jecp.2015.06.003>