Credibility of Stories about Design History

Sergio E. Chaigneau (sergio.chaigneau@uai.cl)
Escuela de Psicología, Universidad Adolfo Ibáñez, Diagonal Las Torres 2640
Peñalolén, Santiago, Chile

Cristián Coo (cristian.coo2005@alumnos.uai.cl)
Escuela de Psicología, Universidad Adolfo Ibáñez, Diagonal Las Torres 2640
Peñalolén, Santiago, Chile

Vicente Soto (vicente.soto2005@alumnos.uai.cl)
Escuela de Psicología, Universidad Adolfo Ibáñez, Diagonal Las Torres 2640
Peñalolén, Santiago, Chile

Abstract

There is evidence that a story about design history is credible to the extent that it coheres with object affordances. Results that support this theory were generally obtained with artificial materials learned in laboratory experiments. In the current experiment, we extend these findings to real artifacts that occur naturally outside the laboratory. We presented participants of different levels of expertise, with real artifacts that cohered to different degrees with a proposed design history, and participants rated either the artifacts' efficiency for the proposed design history function, or the credibility of the received story about their design history. Our results showed that coherent design histories are more credible, that expertise increases the ability to judge efficiency (not surprisingly), but that expertise does not affect the ability to judge credibility to the same extent that it affects efficiency. A post hoc explanation is offered for this interaction.

Keywords: Artifact kinds; design history; explanation; coherence.

Introduction

Knowledge about how objects were created has an important effect on how those objects are conceptualized. For artifact kinds, objects' design histories can become conceptually central. Children and adults prefer to assign artifacts their function according to their design history rather than according to an alternative afforded function (Kelemen, 1999; Rips, 1989), use design history to guide naming (Jaswal, 2006; Matan & Carey, 2001), and can be primed with it during problem solving (Defeyter & German, 2003). The conceptual relevance of design history extends to natural kinds, provided conceptualizers believe these kinds were created by a supernatural being. In one study (Evans, 2001), children's belief in God as the creator of categories was found to correlate with their tendency to reason essentialistically about the nature of animals (i.e., an endorsement of stable category membership). In another study (Diesendruck & Haber, 2009), where children of religious and non-religious environments were compared, it was found that belief in God as creator correlated with an essentialist mode of thinking about social categories (e.g., race, gender, ethnicity), and with a teleological mode of thinking about animals (i.e., a belief that they were made for something).

Similarly to what happens for natural kinds, for design history to have an effect on artifact concepts, conceptualizers need to believe it. It is easy to overlook the fact that knowledge about design history is generally learned from an indirect source. People that know a telephone is used to communicate with other people over long distances, know this because they can experience it in the present. People that know the story about how Bell invented the telephone, know this because they learned it from somebody else, and not because they directly perceived it. In this sense, artifacts are no different from natural kinds created by a supernatural being. For design history to have a conceptual effect, people need to judge it is credible.

Many factors could enter into credibility judgments. Among other factors, a conceptualizer could judge credibility according to the source's trustworthiness, according the status of the putative designer, or according to whether the known object affordances are coherent with the putative design history (cf., Thagard, 2005). Of the factors just mentioned, coherence is the one that has received the most attention in cognitive psychology (e.g., McNamara, Kintsh, Songer, & Kintsh, 1996; Murphy & Medin, 1985; Rehder & Kim, 2006). According to a coherentist view, people have a preference for good or coherent explanations because they approximate truth (see Thagard, 1992, 2007). Therefore, a design history should be credible to the extent that it is coherent (see Thagard, 1989, 2005). In fact, experiments with artifact categories offer confirming evidence. A design history story becomes conceptually central if it offers a plausible explanation for the artifact's properties and function. More specifically, when children receive artifacts with salient properties that are not coherent with the object's putative design history, then that design history does not have conceptual effects (Kemler Nelson, Frankenfield, Morris & Blair, 2000, Exp. 1) and promotes more inquiries signaling conceptual dissatisfaction (Asher & Kemler Nelson, 2008). Because prior work on this topic was done with artificial stimuli in the laboratory, a remaining issue is whether coherence effects on credibility extend to naturally occurring artifact categories. In the current experiment we tested this hypothesis using real objects that people learned about in their everyday dealings (i.e., not in the laboratory).

Another issue that has not been explored up to date is whether expertise has an effect on the credibility of stories about design history. There are reasons to think it could. In the limiting case of a person who knows close to nothing about a given category, then she may be willing to accept a wide range of design histories. Provided these stories are not incoherent with other things she knows, the design history story may only need a loose connection with what she knows about the category to be accepted as credible. A child learning about computers is a case in point. In the opposite limiting case of someone who is an expert in a given domain, a purported design history may need a greater amount of detail to be accepted as coherent. An archeologist examining ancient artifacts is a case in point. This is an important question because if a small amount of knowledge is sufficient to lend credibility to a design history story, then a child or a novice learning a new category may use the design history story to guide the new category's construction.

Experiment's Overview

We presented participants with photographs of four real objects (climbing gear) that they were acquainted with (participants practiced rock climbing), along with a putative design history story that specified ideal conditions of use (always the same for all four objects). Thus, the design history provided an ideal, relative to which the four objects' efficiencies could be ordered. Participants were of different levels of expertise (novice and expert climbers). The four objects were selected by the experimenters so that they could be ordered in decreasing efficiency (three objects were efficient to different degrees, and one was inefficient). Participants were then asked to rate (between participants) either whether they agreed that the artifact was efficient for the stated function (efficiency rating), or whether they believed the artifact's design history story was true (credibility rating).

Predictions

We used three kinds of climbing gear: pulleys, ascenders and belays. Though these are relatively simple mechanical devices, understanding their functions requires experience with the objects themselves and with the conditions in which they can be used. For example, two of our pulleys look very similar to each other, and can be distinguished only based on technical criteria such as weight and ease of use, while a third pulley looks different but can be made to work provided the necessary experience. To understand an object's function, requires knowing not only the object's physical structure, but also the actions that the object affords, the setting in which it operates, and the outcomes that can be expected to occur (cf., Chaigneau, Barsalou & Zamani, 2009). It is likely that our experts would have fuller and more detailed models of the situations referred to by our texts and photographs (cf., Groen & Patel, 1988), and because of this we predicted that experts would have more differentiated efficiency judgments than novices.

Note that in this design, efficiency necessarily covaries with coherence (i.e., efficiency was our proxy for coherence). Because the design history story was fixed (i.e., the same for the four objects), less efficiency for the stated historical function implies that the design history is also less coherent (i.e., it is unlikely that a designer will create an artifact to function with low efficiency for its purported function). Thus, coherence predicts that credibility ratings should pattern similarly to efficiency ratings. (A different way to manipulate coherence is to directly provide participants with knowledge varying in coherence. However, this is possible only when materials are completely artificial, and not when real objects are used.)

Finally, our experiment also informed us of effects of expertise on credibility. Differences between experts and novices would tell us if the amount of knowledge possessed by the conceptualizer affects the credibility of a design history story.

Method

Design and Participants

Participants were 42 vertical rock climbing practitioners who volunteered to participate by login into the experiment through a link placed in various sport related websites. The experiment used a 2 (*rating*: efficiency, credibility) x 3 (*object kind*: pulley, ascender, belay) x 4 (*object efficiency* level) mixed design, with repeated measures on the last factor. The link randomly directed participants to one of the 2 x 3 = 6 between participants cells of the design. Before beginning the experiment, participants informed the difficulty level of their last successful climb. Most participants informed difficulty by using the Yosemite Decimal System (YDS). When a different system was used (British, French), the score was transformed to YDS using a conventional conversion table. The median YDS score was used to divide participants in two expertise levels. This allowed one additional between participants factor in the design (*expertise*: novice, expert). Because of the online nature of the procedure, we continued collecting data until we had a sufficient number of participants in each level of the rating factor (21 for efficiency, 21 for credibility), and about an equal number of participants in each cell of the design (efficiency: 12 novices and 9 experts; credibility: 11 novices and 10 experts). Data from 7 participants were discarded because they did not finish the experiment.

Materials and Procedures

We selected three kinds of climbing gear: pulleys, ascenders and belays. For each kind, three photographs were selected from the Petzl® company website (see Figure 1). Each triad was selected spanning a wide range of efficiencies. To define efficiency, we imagined a novice climber, so that the most efficient object of each kind would be the one that was easiest to use and offered the greatest safety. For example, the a priori most efficient belay for a novice had an assisted breaking mechanism, while the other two belays required increasingly good rope handling. To each triad, an object from a different kind was added to have an inefficient fourth object (e.g., the fourth object for the belay group was really an emergency ascender). Object selection and efficiency estimates were performed by the second and third authors, who practice rock climbing regularly. For each group of four objects, a fictitious design history story was created, stating that each object was designed to be used by novice climbers in some specific situation (see Figure 1 for an example). (Note that we are not tapping here on objective efficiencies, because these artifacts are efficient only relative to conditions of use such that, e.g., a climber that privileges object weight and multifunctionality would probably give a different efficiency ordering.)

Participants started the experiment by clicking on a link that randomly directed them to one of six versions of the experiment, which resulted from combining the 3 kinds of climbing gear with the 2 levels of the rating factor. Thus, each participant rated only one set of four objects with different efficiencies, either in the efficiency rating or in the credibility rating condition (for examples of questions, see Figure 1). For each group of four objects, we created a different pseudo-random sequence of levels of efficiency, taking care that the four objects never appeared in continuous decreasing or increasing efficiency order. Ratings were collected on a 7-point likert scale. Participants could abandon the experiment at any point.

Results

Data were submitted to a 2 (*expertise*: novice, expert) x 2 (*rating*: credibility, efficiency) x 4 (*object efficiency*) mixed ANOVA, with repeated measures on the last factor. (Object kind was treated as a control variable, and was not included as a factor in the analyses.) The overall ANOVA showed a main effect of object efficiency ($F(3, 114) = 40.89$, $MSe = 3.47$, $p < .001$, $R^2 = .52$, power = 1), no effect of expertise ($F(1, 38) = 2.80$, $MSe = 1.66$, $p < .25$, $R^2 = .07$, power = .37), and no effect of rating ($F<1$). The only significant two-way interaction was rating by expertise ($F(1, 38) = 4.19$, $MSe = 1.66$, $p < .05$, $R^2 = .10$, power = .51), but there was no object efficiency by rating interaction ($F(3, 114) = 1.96$, $MSe = 3.47$, $p < .25$, $R^2 = .05$, power = .40), and no efficiency by expertise interaction ($F(3, 114) = 2.0$, $MSe = 3.47$, $p < .25$, $R^2 = .05$, power = .42). Finally, there was a significant three way interaction between expertise, rating and object efficiency ($F(3, 114) = 4.42$, $MSe = 3.47$, $p < .05$, $R^2 = .10$, power = .77).

As Figure 2 illustrates, contrasts on the repeated measures object efficiency main effect showed ratings decreased linearly along a priori object efficiency levels. Level 1 produced significantly greater ratings than level 2 ($F(1, 38) = 10.02$, $MSe = 3.07$, $p < .01$, $R^2 = .21$, power = .87). Level 2 produced significantly greater ratings than level 3 ($F(1, 38) = 8.47$, $MSe = 1.88$, $p < .01$, $R^2 = .18$, power = .81). Level 3 produced significantly greater ratings than level 4 ($F(1, 38) = 27.55$, $MSe = 7.03$, $p < .001$, $R^2 = .42$, power = 1).

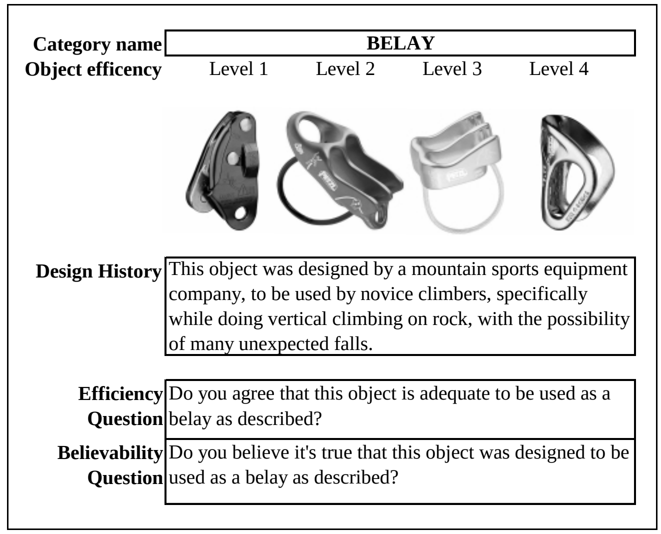

Figure 1: Four objects from the *belay* category, ordered in a priori decreasing efficiency levels. Participants, however, never received objects in this or in the inverse order. Participants saw objects (color photographs from the Petzl® company website) one at a time, each time with the same proposed design history. For each object, participants answered either the efficiency or the credibility question, using a 7-point likert rating scale.

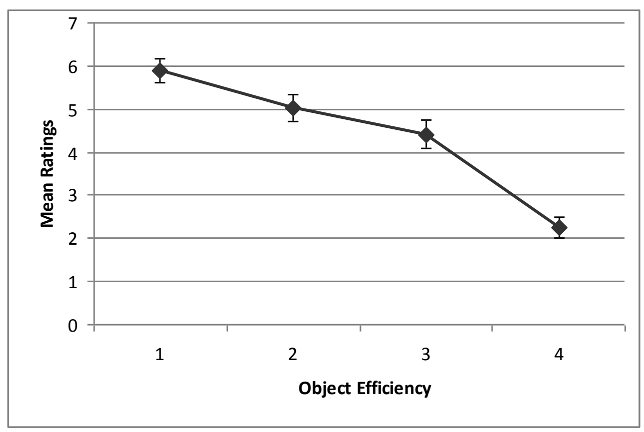

Figure 2: Mean ratings across rating type and object kind. Values in the X axis represent increasingly inefficient artifacts. All differences among adjacent means are significant. Error bars are standard errors.

The significant three-way interaction was explored by simple contrasts between experts and novices for efficiency and credibility ratings, at each level of the object efficiency factor. As Figure 3a shows, efficiency ratings were greater for experts than novices, at all levels of object efficiency except for the inefficient object. At the first, second, and third levels of object efficiency, efficiency ratings were greater for experts than for novices (respectively, $F(1, 38) = 11.46$, $MSe = 2.93$, $p < .01$; $F(1, 38) = 6.51$, $MSe = 3.86$, $p < .05$; $F(1, 38) = 5.46$, $MSe = 4.66$, $p < .05$). At the last level of object efficiency, efficiency ratings for experts were not significantly different from those of novices ($F(1, 38) = 1.89$, $MSe = 2.61$, $p < .25$). In contrast, as Figure 3b shows, when credibility ratings were considered, none of the comparisons between experts and novices at levels of object efficiency were significant (for all levels, $F < 1$).

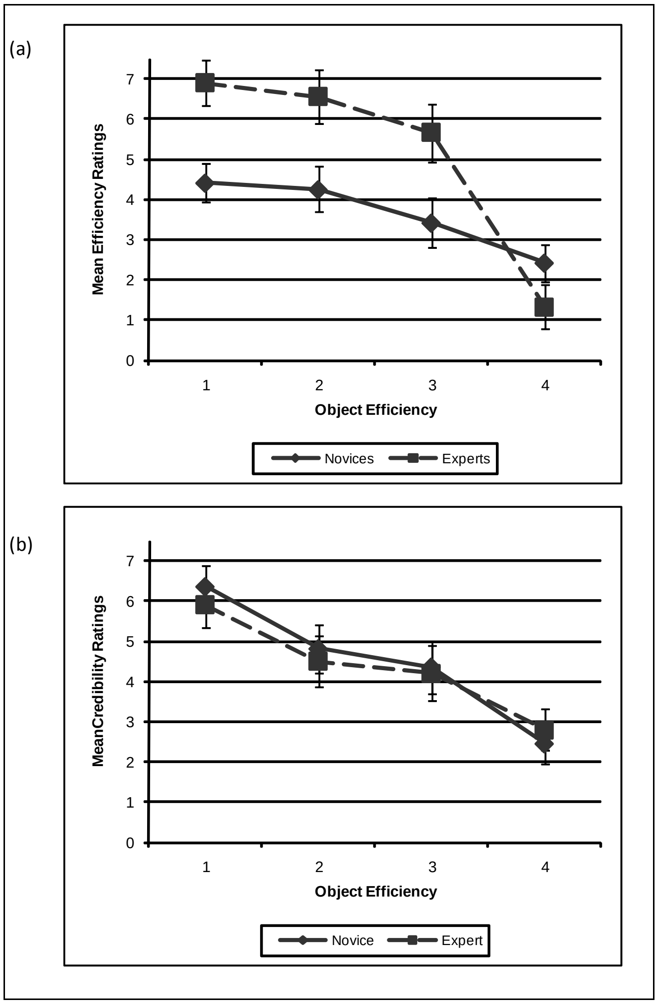

Figure 3: In Panel (a), mean efficiency ratings comparing experts and novices. In Panel (b), mean credibility ratings also comparing experts and novices. The graphs show a three-way interaction between expertise, rating type and object efficiency. Error bars are standard errors.

Discussion

As Figure 2 shows, we successfully graduated levels of efficiency. Our participants rated efficiencies consistently with our a priori judgment, with the four levels showing an almost linear trend of decreasing efficiency. The main effect of the object efficiency factor, suggest this was a powerful source of information in the experiment.

Recall (from the Predictions section) that in our design, object efficiency correlates with design history's coherence, and that therefore coherence predicts credibility ratings would pattern similarly to efficiency ratings. As Figure 3 shows, credibility ratings for experts and novices show a similar linear trend to efficiency ratings. In fact, we found no interaction of a priori object efficiency with rating or with expertise.

The simple contrasts showed that, as predicted, expertise affected efficiency ratings. Novices gave lower ratings overall, and were less able to discriminate efficiencies, offering ratings close to the middle of the scale (see Figure 3a). However, expertise did not affect credibility ratings. Novices judged credibility almost identical to experts, both in pattern and slope (see Figure 3b). We think this is not an issue of lack of power or of the particular materials we used, because our design was sensitive enough to detect differences between novices and experts' efficiency ratings, and yet did not detect differences between novices and experts' credibility ratings.

Regarding things that were created (natural or artificial), our results provide support for the theory that design history is conceptually central if it coheres with objects' affordances. The better an object affords its purported historical function, the more credible its design history story is. Our results extend this theory by showing that the phenomenon happens not only in laboratory experiments, but also in naturally occurring artifact categories.

Our results also draw attention to an interesting difference between the relation of expertise to efficiency, and of expertise to credibility. Though experts and novices differed in their efficiency ratings, with the former providing overall higher and better differentiated ratings, they did not differ in their credibility ratings (at least with our materials and in the range of differences we studied).

We acknowledge there are many interpretations for our three-way interaction, buy we nonetheless want to offer one that seems particularly interesting for setting up more precise experiments. An interpretation for this pattern of results regarding expertise is that the credibility of a design history depends on information that overlaps with information used to compute efficiency but is distinct from it. An example will help to illustrate what we think is happening. Imagine someone who is learning about the *bread knife* category. This category is characterized by two properties: the blade of bread knifes is long and serrated. Imagine, furthermore, that because this individual's experience with bread is limited, she knows about long blades but not about serrated blades. Now, if presented with a set of knifes having different combinations of long or short and straight or serrated blades, and asked to judge their relative efficiencies, she will probably group them in fewer groups and be less discriminating (equivalent to novices in our experiment) than someone with more experience who knows about the two properties (equivalent to experts in our experiment).

However, the question of design history credibility may be solved by appeal to other types of information. Imagine that our subject is told that bread knifes were made to prevent people from crushing slices when cutting a loaf of bread, and then asked to judge if that story about design is credible for each knife in the set. Of course, the knifes' relative efficiencies are still relevant to decide if the design history is coherent and credible (i.e., designers don't create inefficient artifacts), but other types of information may also be useful. The availability of contrast categories that could achieve the same function may be one of them. If our subject can't think of a different category of utensils that can be used to cut bread with similar efficiency than bread knifes, then that may increase the coherence of the story about design history (i.e., artifacts are created to achieve a goal that is typically no achievable as efficiently by other means).

Under this interpretation, our results may show that though novices had less detailed information about the affordances of pulleys, ascenders and belays, they shared with experts other types of information that were relevant for their credibility judgments. Perhaps it was this shared information that made novice's and expert's credibility judgments less distinguishable than their efficiency judgments.

Acknowledgments

We are grateful for comments provided by three anonymous reviewers, which contributed to increase this manuscript's coherence. This work was supported by FONDECYT grant 1100426 to the first author.

References

Asher, Y. M., & Kemler Nelson, D. G. (2008). Was it designed to do that? Children's focus on intended function in their conceptualization of artifacts. *Cognition, 106*(1), 474–483.

Chaigneau, S. E., Barsalou, L. W., & Zamani, M. (2009). Situational information contributes to object categorization and inference. *Acta Psychologica, 130*(1), 81–94.

Chaigneau, S. E., Castillo, R. D. & Martínez, L. (2008). Creators' intentions bias judgments of function independently from causal inferences. *Cognition, 109*, 123–132.

Defeyter, M. A., & German, T. P. (2003). Acquiring an understanding of design: evidence from children's insight problem solving. *Cognition, 89*(2), 133–155.

Diesendruck, G., & Haber, L. (2009). God's categories: the effect of religiosity on children's teleological and essentialist beliefs about categories. *Cognition, 110*, 100–114.

Groen, G. J., & Patel, V. L. (1988). The relationship between comprehension and reasoning in medical expertise. In M. T. H. Chi, R. Glaser, and M. J. Farr (Eds.) *The Nature of Expertise*. Hillsdale, NJ: Erlbaum.

Evans, M. E. (2001). Cognitive and contextual factors in the emergence of diverse belief systems: creation versus evolution. *Cognitive Psychology, 42*, 217–266.

Jaswal, V. K. (2006). Preschoolers favor the creator's label when reasoning about an artifact's function. *Cognition, 99*, B83–B92.

Kelemen, D. (1999). The scope of teleological thinking in preschool children. *Cognition, 70*, 241–272.

Kemler Nelson, D. G., Frankenfield, A., Morris, C., & Blair, E. (2000). Young children's use of functional information to categorize artifacts: Three factors that matter. *Cognition, 77*, 133–68.

Matan, A., & Carey, S. (2001). Developmental changes within the core of artifact concepts. *Cognition, 78*, 1–26.

McNamara, D.S., Kintsh, E., Songer, N., & Kintsh, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, *14*(1), 1–43.

Murphy, G.L., & Medin, D.L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92*(3), 289–316

Rehder, B. & Kim, S. (2006). How causal knowledge affects classification: A generative theory of categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 659–683.

Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and Analogical Reasoning*. New York: Cambridge University Press.

Thagard, P. (1989). Explanatory Coherence. *Behavioral and Brain Sciences, 12*, 435–502.

Thagard, P. (1992). Conceptual Revolutions. Princeton, NJ: Princeton University Press.

Thagard, P. (2005). Testimony, credibility, and explanatory coherence. *Erkenntnis, 63*, 295–316.

Thagard, P. (2007). Coherence, truth, and the development of scientific knowledge. *Philosophy of Science,74*, 28–47.