Title
Is the Centrality of Design History Function an Effect of Causal Knowledge?

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

ISSN
1069-7977

Authors
Puebla-Ramirez, Guillermo
Chaigneau, Sergio

Publication Date
2011

Peer reviewed

Is the Centrality of Design History Function an Effect of Causal Knowledge?

Guillermo Puebla-Ramírez (pueblaramirezg@gmail.com)
Unidad de Análisis e Investigación Institucional, Universidad de Tarapacá, General Velásquez 1775
Arica, Chile

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

Abstract

Design history function (i.e., what an artifact was made for) is a central aspect of artifact conceptualization. A generally accepted explanation is that design history is central because it is the root cause for many other artifact properties. In Exp. 1, an inference task allowed us to probe participants' causal models, and then to use them when making predictions for Exp. 2. Design history was, in fact, part of what participants viewed as conceptually relevant. Predictions for Exp. 2 were derived using the currently most comprehensive theory about how causal knowledge affects categorization. Our results show that though participants used design history, functional outcome and physical structure to conceptualize artifacts, the effect of design history was independent from knowledge of physical structure and functional outcome. This result is inconsistent with a causal knowledge explanation of design history's conceptual centrality.

Keywords: categorization; causal reasoning; essentialism; artifacts.

Introduction

Imagine you inherited an antique sewing machine. It comes in a beautiful cabinet, so you decided to use it in your living room as a table to display stuff. Imagine now that a guest comments about the beautiful table. A good bet is that your reaction would be to tell the visitor that the object is really a sewing machine cabinet, but that you currently use it as a table. The general phenomenon illustrated here is that there is a preference to conceptualize artifacts according to what they were designed for (their *design history function*) rather than according to an alternative but current function. First described by Lance Rips (1989), this is a robust phenomenon, valid across different paradigms (e.g., Chaigneau, Castillo & Martínez, 2008; Gelman & Bloom, 2000; Defeyter, Avons, & German, 2007; Defeyter & German, 2003; Jaswal, 2006; Matan & Carey, 2001), age levels (Gutheil, Bloom, Valderrama, & Freedman, 2004) and cultures (German & Barrett, 2005).

The favored explanation for this phenomenon is that it occurs because people view design history as the essence of artifacts (e.g., Bloom, 1996, 1998, 2007). Medin and Ortony's (1989) *psychological essentialism*, holds that when someone categorizes objects, she focuses on what she knows (cognitively or metacognitively) about the cause of the object's apparent properties, more than she focuses on the apparent properties themselves. An essence, in this view, is an often unobserved root cause that explains many of an entity's surface features (Ahn, Kalish, Gelman, Medin, Luhman, Atran, Coley, & Shafto, 2001). Correspondingly, because the design history function can be reasonably viewed as the root cause of an artifact's physical structure and use, several authors have assumed that this is why people use design history function (and not current function, nor object appearance) for conceptualization (e.g., Matan, & Carey, 2001; Kemler-Nelson, Frankenfield, Morris, & Blair, 2000; Asher & Kemler Nelson, 2008). Importantly, in this view the relevance of the design history is a consequence of people's causal knowledge about artifacts.

In the experiments we report here, we tested if the influence of the design history function is a case of causal-based categorization. If the centrality of design history is a consequence of people's causal knowledge, then its influence on category membership judgments should be consistent with documented effects of causal knowledge on categorization. Of particular concern for us is the *causal status effect* phenomenon (Ahn, Gelman, Amsterlaw, Hohenstein and Kalish, 2000; Ahn, Kim, Lassaline, & Dennis, 2000; Meunier & Cordier, 2009; Rehder & Kim, 2010), in which causes are more important that their effects. To make predictions about the design history's causal influence, we draw heavily on Rehder and collaborators' work about the influence of causal knowledge on categorization (i.e., the generative model; Rehder, 2003a, 2003b, 2010; Rehder & Kim, 2006, 2010). Our aim is not to test the generative model, but because this theory accounts for many different phenomena on causal categorization, our aim is to use it as a benchmark to assess if the conceptual relevance of the design history function can be explained as a causal-based categorization phenomenon. If the conceptual centrality of design history function is explained by causal knowledge, people that use design history function to categorize should show telltale signs of causal categorization.

Causal Influences on Category Judgments

Rehder's research program identifies two routes for causal knowledge's influence on categorization (Rehder, 2010). In *the explicit route*, people treat observed properties as evidence of unobserved properties, and then use these inferred properties for categorization judgments. These inferences can be retrospective (e.g., knowing that A $\rightarrow$ B, using the known presence of B to infer the presence of the unobserved A, and then using this inferred A to categorize;

Rehder & Kim, 2009) or prospective (e.g., using the known A to infer the presence of B; Rehder, 2007; Chaigneau, Barsalou, & Sloman, 2004). Our current Exp. 1 used a prospective reasoning task, allowing us to determine which information participants used to make their inferences.

In contrast, in *the implicit route*, people estimate whether a configuration of known properties (i.e., an exemplar) could be generated by the category's implicit causal model. For example, if a category's implicit causal model is A → B → C, and if links are probabilistic, each successive property is generated with less certainty, and a causal status effect obtains (i.e., A is conceptually more central than B, and B than C). In contrast, if links are deterministic, then all properties are equally certain and no causal status effect obtains (i.e., A, B and C are equally central). In the implicit route, not only individual properties matter for categorization, but also combinations of properties. Simply put, if two properties are causally linked, then they should be correlated (i.e., if one is present/absent, so is the other). For example, if people believe that having large wings causes birds to fly, then an animal that has small wings an flies is a poorer category member than an animal that has small wings and does not fly. These interactions among properties have been found to have larger effect sizes than the effects of individual properties in categorization judgments (reviewed in Rehder, 2010). Our current Exp. 2 used an implicit reasoning task, allowing us to assess if causal reasoning could account for our data.

Experiments' Overview

In the current experiments, participants were presented with scenarios describing a novel artifact's design history (H), its physical structure (P), an agent's goal when using the artifact (G), and the agent's action (A) (and the functional outcome (O), but only in Exp. 2), and asked to rate its category membership. Exp. 1 used an explicit causal reasoning task. Participants were provided with information about H, P, G and A (but not O), and we predicted that they would use the observed properties to infer the state of the unobserved property O, and then use that inferred property to categorize. Results from this experiment allowed us to determine which information participants used for their inferences, and also to hone in on the implicit causal model they used. In Exp. 2, we used an implicit causal reasoning task. Participants were provided with information about H, P, G, A and O, and asked to rate category membership. Given our assessment of participants' implicit causal models in Exp. 1, the generative theory makes clear predictions about the relative weights of properties for category membership ratings. Comparing our obtained pattern of results with theoretical predictions, allowed us to appraise whether participants were doing causal reasoning or not.

On both experiments, we analyzed ratings using Rehder's regression method (2003a, 2003b, 2010). In this method, participants provide category membership ratings for all possible property combinations, allowing the computation of individualized regression equations. Participants in our experiments were presented with a category with causal knowledge regarding 5 binary valued properties. H could describe the artifact being designed towards functionality x or functionality y. P could be described as adequate to achieve functionality x, or not adequate to achieve it. G could be described as intentional and coherent with functionality x, or accidental and not coherent with functionality x. A could be described as coherent with functionality x, or not coherent with functionality x. Finally (but only on Exp. 2) O could be described as achieving or not functionality x. In consequence, participants in Exp. 1 rated 2^4=16 scenarios, and participants in Exp. 2 rated 2^5=32 scenarios. The baseline scenario (i.e., all components coherent with functionality x) was always rated first. Because each participant provided 2^n data points for each variable, a regression equation for each participant was computed, with H, P, G, and A (plus O in Exp. 2) as predictors, and rating as criterion. Regression coefficients for each participant were then used as individual data points reflecting the contribution of each predictor variable to the ratings.

Experiment 1

In this experiment, participants were provided with information about H, G, A and P. No information was provided about O. Because O is arguably the end node of an artifact's causal model, we predicted that participants would engage in explicit causal reasoning to infer O given the known information, and then use the inferred O to categorize. From prior studies, we assumed that the causal model participants would use was H → P →O ← A ← G (see Figure 1a). This is the causal model obtained for scenarios similar to the present ones in Chaigneau, Barsalou & Sloman (2004; hereafter referred to as CB&S).

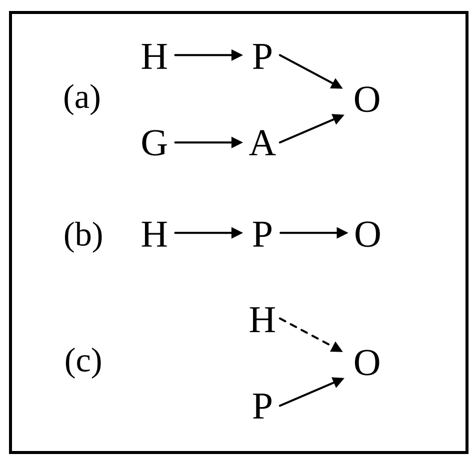


Figure 1: Panel (a) shows the predicted causal model for scenarios in Exp. 1. Panels (b) and (c) show two possible causal models used by our participants, based on results for Exp. 1. The dotted line in panel (c), reflects a weak causal link from H to O.

Because bayesian models (of which the generative model is one) predict that when P and A are specified (the proximal causes), they determine the state of O independently from the state of H and G (the distal causes), we predicted that P and A would show high regression coefficients, while H and G would show significantly lower ones (this was also one of the main results in CB&S). In other words, we expected participants to respect the Markov condition in causal reasoning (Hausman & Woodward, 1999). Additionally, regression weights would inform us which properties participants used for their judgments.

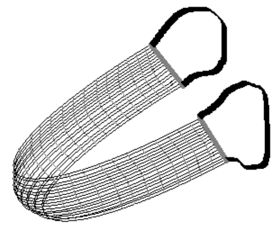
Method

Design and Participants Twenty-four Adolfo Ibáñez undergraduates participated in this study (native Spanish-speakers). Participants were randomly assigned to one of 3 artifacts and one of 4 pseudo-random order of scenarios.

Materials Three novel artifact categories were tested ("peinador", "cazador de peces" and "tatuador"; respectively, "hair-brusher", "fish-catcher" and "tattoo-maker") and 16 scenarios for each category. Each category was designed to afford two plausible functions, one serving as cue to name the artifact. For example, the fish-catcher consisted of a net of vegetable fibers which could (in principle) be used both to catch fish or to carry stones. The cue function was fixed across all scenarios so the question was always the same (e.g., Would you say that this object is a fish-catcher?). Scenarios described one character that created an object and a second character that used it. A graphic depiction of the artifact's physical structure was included in all scenarios. As an example, Figure 2 shows the fish-catcher scenario specifying all elements as adequate (i.e., baseline). When H was compromised, the designer created the object for one function, but the second character used it for a different function (e.g., a net designed to carry stones which is then used as a fish-catcher). When P was compromised, the artifact's physical structure was described and depicted as not affording its cue function (e.g. a net with several holes on it). When G was compromised, the second character's actions were described as accidental (e.g., the second character performed the appropriate actions but was playing and not intending to catch fish). When A was compromised, the second character was described as not performing the appropriate actions (e.g., shaking the net just under the water's surface instead of keeping is stretched and still). Thus, the 16 scenarios for each category presented participants with all combinations of adequate and compromised H, G, A and P.

Procedure Initially, participants received the instructions in writing but also heard them read aloud by the experimenter. Later, participants worked individually. Participants received two training scenarios, which described the creation and use of a hammer. One of these scenarios was a baseline (i.e., all properties adequate), and the second scenario presented the opposite extreme of the scale (i.e., all components compromised). Ratings were performed on a 7-point scale, with 1 always reflecting the low-end ("no") and 7 the high-end ("yes") of the scale.

In an ancient culture, a settler called Kne-Mû wanted to make an object to catch small fish living in large numbers in certain streams. Because he didn't have an object to do that, he decided to make it. The object consisted of a series of intertwined vegetable fibers. On each side, the object had handles (as shown in the picture).



One day, another settler called Knat-knê wanted to catch some small fish from a stream. He found the object Kne-Mû made and thought that it would be useful for catching fish. Knat-knê grasped the object by both handles and kept it stretched just below the stream's surface.

Question: Would you say that this object is a fish-catcher?

Figure 2: Baseline fish-catcher scenario in Exp. 1. In Exp. 2, the scenario also provided information about the event's outcome, by adding: "As a result of the events described, fish in the stream were trapped in the vegetable fibers.")

Results To determine the importance of properties we analyzed participants' ratings by performing a multiple regression for each participant. Four predictor variables were coded as -1 if the feature was compromised and +1 if it was adequate. The regression weight associated with each predictor represents the influence that each element had on ratings. Additionally 6 predictor variables represented the two-way interactions between the four elements. Each of these was coded as -1 if a pair of elements had distinct values and +1 if they had the same value. Note that in this method of analysis, participants provide category membership ratings for all possible property combinations, allowing the computation of individualized regression coefficients, but statistical tests are performed considering the coefficients' variance across participants (i.e., not the significance of the individual coefficients).

Preliminary analyses showed that regression weights for the 6 interaction terms were not significant. Because of this, the following analyses consider only the individual terms. Averaged regression weights over participants for H, G, A, and P are presented in Figure 3. There were no effects of which object participants rated, or of which of the 4 pseudo-random scenario orders participants received, and thus results were collapsed over these factors. To test the differences between the regression weights given to H, G, A and P, an ANOVA with repeated measures was conducted with *individual terms* (4 levels: H, G, A, P) as the single factor. A violation of the sphericity assumption was handled by correcting degrees of freedom with Huynh-Feldt's epsilon. Sphericity was addressed likewise in Exp. 2. For clarity of presentation, degrees of freedom are presented without adjustment here and elsewhere. There was a main effect of individual terms ($F(3, 69) = 9.61$, $MSe = .471$, $p < .001$, $R^2 = .30$, power = 1).

Post hoc tests on the repeated measures factor (with Bonferroni adjustment), revealed that the regression weight associated with P was significantly greater than H, G and A (all *p*s < .05) and that the regression weight associated to H did not differ from those of G or A (both *p*s > .05). Finally, *t* tests showed that only the regression weights for H and P were significantly different from zero ($t(23) = 2.78$, $p < .05$; $t(23) = 4.98$, $p < .001$, respectively).

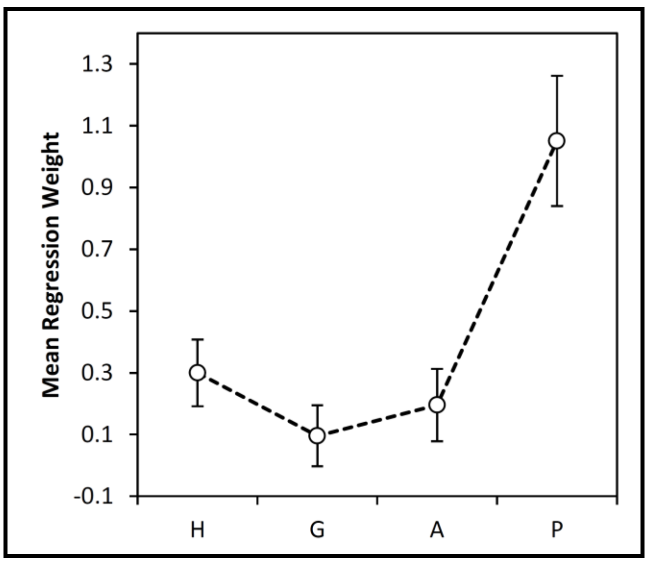


Figure 3: In Exp. 1, mean regression weights for history (H), agent goal (G), agent action (A) and physical structure (P). Only P and H coefficients were significantly greater than zero. Bars are standard errors.

Discussion

Results suggest that participants did not use model 1a, because G and A did not influence their ratings. Because only P and H were significantly different from zero, taken as a group, participants appear to have used a model similar to 1b or 1c. The causal Markov condition predicts that in a chain model like 1b, the distal cause will exert less influence on the outcome than the proximal cause. Consistently with this prediction, results showed that the coefficient for P was greater than the coefficient for H (consistently with results in CB&S). However, model 1c could also account for these results. Lombrozo (2010) has proposed people can treat human intentions (e.g., the designer's intention) as metaphorical mechanisms of causal transmission. Assuming that a metaphorical cause (H → O) has lower strength than a mechanical one (P → O), model 1c could explain why P and H affected ratings, but H had a weaker effect. Models 1b and 1c were used to generate predictions for Exp. 2.

Experiment 2

Exp. 2 assessed the importance of H, G, A, P and O on categorization judgments, now with an implicit causal reasoning task. Because participants' ratings in Exp. 1 were not influenced by G nor A, we predicted that in Exp. 2 G and A would not show significant regression coefficients. Considering the model in Figure 1b, the generative model theory predicts that if participants interpret causal links as deterministic, the coefficients in the implicit reasoning task will be H = P = O. On the other hand, if participants interpret causal links as probabilistic, the theory predicts regression weights H > P > O (i.e., a causal status effect). Considering the model in Figure 1c, and given that it has a weak causal link from H to O, the generative model theory predicts that participants should weigh less deviations from H's baseline value than from P's baseline value. This prediction is derived because the weak causal link implies a small correlation between H and O, and therefore deviations from H's baseline value should have a lesser impact on ratings than deviations for P.

Method

Design and participants Thirty Adolfo Ibáñez and Tarapacá University undergraduates participated in this study (native Spanish-speakers). Participants were randomly assigned to one of 3 artifacts and one of 5 pseudo-random order of scenarios.

Materials Materials were the same of Exp. 1, except that information about the functional outcome was systematically manipulated. When O was compromised, the outcome related to the cue function was described as not happening (e.g., for the fish-catcher artifact, fish were not caught in the net). This meant that scenarios had 5 binary properties (H, G, A, P and O), and that participants provided $2^5 = 32$ ratings.

Procedure The procedure was identical to that of Exp. 1.

Results Regression weights for the 5 individual terms and for the 10 two-way interaction terms were computed as described for Exp. 1. Preliminary analyses revealed that regression weights for the interaction terms were not significant. Regression weights averaged over participants for H, G, A, P and O are presented in Figure 4. Again, there were no effects of neither which object participants rated or of the 5 pseudo-random scenario orders, and thus results were collapsed over these factors. An ANOVA with repeated measures was conducted with *individual terms* (5 levels: H, G, A, P and O) as the single factor, which revealed a main effect ($F(4, 116) = 23.35$, $MSe = .30$, $p < .001$, $R^2 = .45$, power = 1).

Post hoc tests were conducted on the repeated measures factor (with Bonferroni adjustment). This analysis showed that the regression weight associated with O was significantly different from G, A and P (all *p*s < .05) but not significantly different from H ($p > .05$). Additionally, P was significantly different from G and A (both *p*s < .05), but not from H ($p > .05$). The regression weight associated with H only differed from those of G and A (both *p*s < .05). Finally, *t* tests showed that only the regression weight of H, P and O were significantly different from zero ($t(29) = 5.91$, $p < .001$; $t(29) = 3.53$, $p < .01$; $t(29) = 7.70$, $p < .001$, respectively).

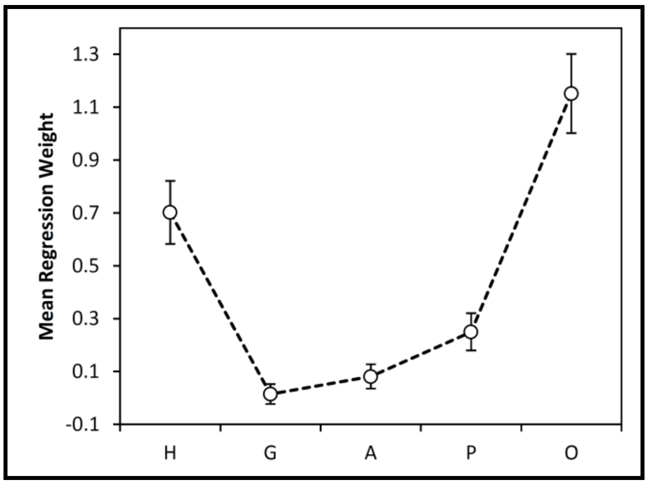


Figure 4: In Exp. 2, mean regression weights for history (H), agent goal (G), agent action (A), physical structure (P) and outcome (O). Only H, P and O were significantly greater than zero. Bars are standard errors.

Finally, we wanted to test if the pattern in Figure 4 resulted from aggregating data from groups of participants who adopted different strategies. It might be that a group of participants decided based on H and did not pay attention to O, while another group decided based on O and did not pay attention to H. If this were true, we should find a negative correlation between H and O coefficients and two distinct groups of data points in the scatterplot (i.e., individuals with high coefficients for H and close to zero for O and vice versa). This was not what data showed. The correlation between H and O turned out to be negative but small and non-significant ($r(28) = -.20$, $p = .29$). Visual inspection of the scatterplot revealed that 3 participants appeared to use the abovementioned strategies, but that a great majority of participants integrated H and O in their categorization judgments and exhibited individual patterns of coefficients similar to the aggregated pattern.

Discussion

As predicted, neither G nor A showed significant coefficients. This lends support to our assumption that participants used the same information to make their judgments in both experiments. Coefficients for O were greater than coefficients for P, with H somewhere in between. The significant difference between O and P, rules out the explanation that participants used model 1b with deterministic links, because this should produce that coefficients H = P = O. At first glance, the relatively high coefficient for H could be interpreted as a causal status effect. However, if participants used chain model 1b and interpreted links as probabilistic (which is a condition that could produce a causal status effect), the curve should show a negative slope, with H > P > O. O's high coefficients speak against this account. Model 1c does not fare better. This model predicts lower coefficients for H than for P, while results showed that H was nominally higher than P. For the sake of completeness, we considered one additional model. Model 1c with two deterministic links could account for our results. This model explains that O has a higher coefficient because, as it has 2 causes, it has a high probability of being generated, while P and H should have about equal weights. However, recall that this last model is not consistent with Exp. 1's results, and so it is also unable to account for the complete pattern of results.

Even further evidence for the absence of implicit causal reasoning in Exp. 2 is that prior research (reviewed in Rehder, 2010) finds that in the implicit causal reasoning task, interactions among properties account for a greater amount of variance than individual properties, while data in Exp. 2 did not show such interactions.

In summary, we find very little evidence that our participants in Exp. 2 did causal reasoning, and yet, H was as conceptually central as O in their ratings. This, we think, shows that design history function can have an important influence on categorization without traces of causal essentialist reasoning in particular, or causal reasoning in general.

General Discussion

As in CB&S, in Exp. 1 participants were not provided with descriptions of O, thus promoting prospective inferences. Consistently with results in CB&S, in Exp. 1 H lost relevance for categorization, presumably because it was partially screened-off by P, which was O's proximal cause. In Exp. 2, when—in contrast to Exp. 1 and to CB&S—information about O was provided, H became at least as important as P for categorization. Contrary to causal essentialism, this increased relevance of H does not correspond with known effects of causal knowledge on categorization.

Simple heuristic processes are unlikely explanations of our results. One alternative is that participants in Exp. 2 categorized based on a simple property count. Our data speaks against this alternative, given that participants consistently used some properties to guide judgments and disregarded others. Another alternative is that participants in Exp. 2 categorized based on diagnostic properties. We think this is not plausible. There is no a-priori reason to think that some properties were more diagnostic than others. Think of a hammer as an example. Using an object with a hammering motion (i.e., A) appears to be at least as diagnostic of the hammer category as is achieving the goal of inserting nails (i.e., O). Also, given that P was the most informative property in Exp. 1, one would expect that it would be at least as diagnostic as O, but this is not what our results in Exp. 2 showed.

In conclusion, based on our participants' response pattern, the current work shows that the conceptual centrality of design history function is not easily explained by causal-based categorization in general, nor by causal essentialism in particular. Our results, especially those of Exp. 2, suggest that design history's contribution to artifact category membership follows an independent mechanism, and is not mediated by causal reasoning about the effect of physical structure on functional outcome.

Acknowledgments

We are grateful to Bob Rehder for his comments on these experiments, and for input on an earlier draft. We also wish to thank the reviewers for pointing out some necessary clarifications. Finally, we wish to thank Cristián Coo, Vicente Soto and Mauricio Ríos for their help in data collection. This work was supported by FONDECYT grant 1100426 to the second author.

References

Ahn, W., Gelman, S. A., Amsterlaw, A., Hohenstein, J., & Kalish, C. W. (2000). Causal status effect in children's categorization. *Cognition, 76*, B35-B43.

Ahn, W., Kalish, C., Gelman, S., Medin, D., Luhman, C., Atran, S., Coley, J. & Shafto, P. (2001). Why essences are essential in the psychology of concepts. *Cognition, 82*, 56–69.

Ahn, W., Kim, N. S., Lassaline, M. E., & Dennis, M. J. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology, 41*, 361-416.

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.

Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition, 60*, 1-29.

Bloom, P. (1998). Theories of artifact categorization. *Cognition, 66*, 87-93.

Bloom, P. (2007). More than words: A reply to Malt and Sloman. *Cognition, 105*, 649–655.

Chaigneau, S.E., Barsalou, L.W., & Sloman, S. (2004). Assessing the causal structure of function. *Journal of Experimental Psychology: General, 133*, 601–625.

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., Avons, S.E., & German, T.C. (2007). Developmental changes in information central to artifact representation: evidence from 'functional fluency' tasks. *Developmental Science*, *10*(5), 538–46.

Defeyter, M.A. & German, T.P. (2003). Acquiring an understanding of design: evidence from children's insight problem solving. *Cognition, 89*(2), 133–55.

Gelman, S.A. & Bloom, P. (2000). Young children are sensitive to how an object was created when deciding what to name it. *Cognition, 76*, 91–103.

German, T.P. & Barrett, H.C. (2005). Functional fixedness in a technologically sparse culture. *Psychological Science, 16*(1), 1–5.

Gutheil, G., Bloom, P., Valderrama, N., & Freedman, R. (2004). The role of historical intuitions in children's and adults' naming of artifacts. *Cognition, 91*(1), 23–42.

Hausman, D. & Woodward, J. (1999). Independence, Invariance, and the Causal Markov Condition. *British Journal for the Philosophy of Science, 50*, 521–83.

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

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*, 133–168.

Lombrozo, T. (2010). Causal-explanatory pluralism: How intentions, functions, and mechanisms influence causal ascriptions. *Cognitive Psychology, 61*(4), 303–332.

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

Medin, D.L., & Ortony, A. (1989). Psychological Essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning*. New York: Cambridge University Press.

Meunier, B, & Cordier, F. (2009).The role of feature type and causal status in 4–5-year-old children's biological categorizations. *Cognitive Development, 24*, 34-48.

Rehder, B. (2003a). Categorization as causal reasoning. *Cognitive Science, 27*(5), 709–748.

Rehder, B. (2003b). A causal-model theory of conceptual representation and categorization. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*(6), 1141–1159.

Rehder, B. (2007). Essentialism as a generative theory of classification. In A. Gopnik, & L. Schultz (Eds.), *Causal learning: Psychology, philosophy, and computation*. Oxford, UK: Oxford University Press.

Rehder, B. (2010). Causal-Based Categorization: A Review. In Brian H. Ross (Ed.), *Psychology of Learning and Motivation, 52*, 39–116.

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.

Rehder, B. & Kim, S. (2009). Classification as diagnostic reasoning. *Memory & Cognition, 37*, 715–729.

Rehder, B. & Kim, S. (2010). Causal status and coherence in causal-based categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*