The (not so) strange death of stimulus equivalence

Dermot Barnes-Holmes
National University of Ireland, Maynooth

Steven C. Hayes
University of Nevada, Reno

Bryan Roche
National University of Ireland, Maynooth

Undertakers play a necessary if somewhat unpopular role in modern life. And this is how we see Tonneau's article — as the undertaker who is kindly helping bury our close friend, stimulus equivalence. We suspect that few will thank Tonneau for providing this unpleasant service, but it needs to be done.

Funerals are not pleasant occasions, but they can serve an important social and psychological function in preparing those left behind to move on. In this case, moving on will be neither simple nor easy. Like a widowed spouse considering the future, equivalence researchers may be frightened by what seems to lie ahead. For some, stimulus equivalence was the behavior-analytic answer to the cognitive challenge. Tonneau has helped show how thin that claim has become. The demise of equivalence may appear to threaten the future of behavior analysis – at least for those who feel (correctly) that providing an adequate account of language and cognition is essential. Hopefully, however, this very threat will galvanize the field into action. Human language and cognition stand before behavior analysis like a vast mountain range. Basic behavioral researchers must not turn back from this challenge if their science is to survive.

In this commentary on Tonneau's article we will echo some of his concerns pertaining to the study of stimulus equivalence. However, we will also explain why we believe that his focus on what he calls functional equivalence and stimulus correlations is problematic. Finally, we will argue that a modern behavioral approach is needed to scale the mountains of human language and cognition, and this approach is fundamentally operant in nature. We will offer Relational Frame Theory as a map for the course ahead (see Hayes, Barnes-Holmes, & Roche, 2001).

Stimulus Equivalence as an Exhausted Research Paradigm

Tonneau questions the relevance of the data gathered by equivalence-class researchers to their stated purpose – the understanding of symbolic behavior. In large part we agree. Stimulus equivalence is simply too narrow and too limited to carry the weight some have put on it. It is a research preparation that shows every sign of exhaustion. The warning signs are myriad, but two stand out.

Limited Applied Impact

Behavior analysis is a field that develops behavioral principles in order to understand human complexity (Skinner, 1938, p. 441), not as an end in itself. After 30 years of research, stimulus equivalence has produced little impact on applied behavior analysis. Only very rarely, for example, do the pages of the *Journal of Applied Behavior Analysis* contain an article bearing on the application of stimulus equivalence. If stimulus equivalence is an adequate account of human symbolic behavior, surely much more should be expected.

Address correspondence concerning this article to E-mail: Dermot.Barnes-Holmes@may.ie or Hayes@SCS.UNR.EDU

Limited Basic Impact

Basic behavior analysis, and the field as a whole, needs a theory that will lead to an experimental analysis of human language and cognition. If such a theory were available, we would expect it to inspire basic research on intelligence, problem solving, emotion, reasoning, language development, and a whole host of similar areas. Stimulus equivalence has failed to do so. Instead of broadening over time, the focus of equivalence research seems to be narrowing to increasingly arcane issues, disconnected from human language and cognition as a general phenomenon. What was once an *outcome* that might orient the field to key process issues in the analysis of language, is now being studied as an end in itself.

Considered as an end in itself, equivalence is of dubious value. It has already been studied for 30 years, and the field could spend another 30 years in research primarily characterized by its precise irrelevancy. Behavior analysis does not have that time to waste.

As a result of this narrowing and the limits of the phenomenon itself, the study of stimulus equivalence seems to be adding little that cognitive psychology had not already considered. The rich literatures on semantic network theory and transitive inference, for example, provide conceptual and empirical analyses that can incorporate the phenomenon of equivalence class formation and a great deal more as well.

We are not calling for behavior analysts to embrace cognitive psychology. To the contrary, it is precisely because it is the mission of behavior analysis to provide its own account of human complexity that this solution is unacceptable. But the field cannot wait forever for real progress. If stimulus equivalence is not the vehicle for fundamental progress - and it seems undeniably clear that it is not - we must move on to approaches that are more productive.

Symmetry and Backward Associative Conditioning

Our foregoing concerns clearly echo some of Tonneau's criticisms of the stimulus equivalence research program. However, we do not share all of his views. The concept of stimulus equivalence has oriented the field toward an important property of human language and cognition, that of the bi-directionality of stimulus relations ("symmetry"). Curiously, this is precisely the area that Tonneau brushes over in his analysis of functional equivalence.

The four types of function transfer that Tonneau describes in his article could be based on varieties of forward associative conditioning (sensory preconditioning, second-order conditioning, etc.), and he argues quite reasonably that transitivity in matching to sample could also be considered in the same light. Then on page 19, *in parentheses*, he suggests that symmetry might also yield to a similar analysis, but he refuses to "press such arguments any further." It is this property of bi-directionality (particularly when it is combined with transitivity), that most requires an account, because it seems to require *backward associative conditioning*. Backward conditioning is a notoriously weak effect that occurs in a rather restricted set of conditions (see Spetch, Wilkie, & Pinel, 1981). Hall (1996), a leader in the field of associative learning, has stated that backward associations in matching to sample contexts "are not readily formed" (p. 248).

Interestingly, Hall (1996) has also argued that when backward conditioning effects apparently occur, using matching to sample, they may actually reflect a mediated form of forward conditioning. Imagine, for example, that a discriminative function transfers from C to B after A-B and A-C pairings. Apparently, a result such as this would require backward associative conditioning (i.e., A may acquire some of the functions of B via forward conditioning, but C can only acquire the functions of A, and thus B, via backward conditioning). Several researchers (Hall, 1996; L. J. Hayes, 1992; Urcuioli, 1996) have argued, however, that C-B transfer, in this case, may be based on a mediated form of forward conditioning (i.e., mediated generalization) in the *absence* of any backward associative processes.

According to this account, A-B pairing establishes a private perceptual function (or representation) of B whenever A is presented (i.e., when the subject sees A, an image or representation of B immediately occurs). Thus, whenever A and C are paired, the private image of B is also paired

with C (i.e., when the subject sees A, s/he immediately thinks of B but then C is presented, and thus the private image of B occurs before C). In this case, the discriminative functions of C may transfer to B via forward respondent conditioning.

Urcuioli (1996) also argued that in order for this type of mediated generalization to occur, the A-B training must occur before the A-C training. If, however, A-B and A-C training occur concurrently, "it is unlikely that [the A] samples could generate the necessary prospective mediators to support transfer" (Urcuioli, 1996, p. 65) to the B stimuli. (Hall also implies this argument in his analysis, but he does not state it explicitly). In many equivalence studies, however, A-B and A-C training trials are presented concurrently (e.g., Barnes-Holmes, Keane, Barnes-Holmes, & Smeets, 2000; Barnes & Keenan, 1993), and thus by this account C-B transfer should *not* have occurred unless one wishes to invoke the problematic concept of backward associative conditioning.

Given these difficulties, it is perhaps not hard to understand why Tonneau refused to address the issue of symmetry in his article. If we are to take his arguments seriously, however, he will have to provide an account of those instances of function transfer that appear to require backward associative conditioning. Without this, his work simply questions the set-theoretic analysis of stimulus equivalence, but offers nothing else in its place. As it is, Tonneau has pointed to stimulus associations or correlations, rather than equivalence class formation, as the basis for function transfer, but he has completely failed to address the very issue that made the concept of stimulus equivalence necessary in the first place — the emergence of bidirectional stimulus relations.

Relational Frame Theory: Moving Beyond Function Transfer

One "solution" to the problem of bidirectionality is simply to accept that backward associative conditioning occurs readily and in most conditions with humans, but with great difficulty in nonhumans. This is not unlike Sidman's argument that equivalence is a basic stimulus function (Sidman, 1994). It "solves" the problem of stimulus equivalence, but as we will discuss below it leaves untouched the substantial evidence that bidirectionality characterizes many different types of stimulus relations (e.g., Dymond & Barnes, 1995, 1996; Roche & Barnes, 1996, 1997; Roche, Barnes-Holmes, Smeets, Barnes-Holmes, & McGeedy, 2000; Steele & Hayes, 1991). Something much more flexible than backward conditioning seems needed.

A ready behavioral alternative is available: the bidirectional stimulus relations observed in equivalence, and other types of stimulus relations, reflects a generalized operant behavioral process (Barnes-Holmes & Barnes-Holmes, 2000; Hayes, Gifford, & Wilson, 1996). This is the core conception of Relational Frame Theory (RFT; Hayes, Barnes-Holmes, & Roche, 2001). Associative or stimulus-correlational effects are certainly also involved (Barnes, 1994; Leader, Barnes, & Smeets, 1996), but they do not play a primary explanatory role in RFT (Barnes & Roche, 1996). In fact, from the RFT perspective, the effects of stimulus correlations on behavior may often depend upon the generalized operant process of relational framing itself (see Barnes, Smeets, & Leader, 1996, pp. 167-168; Leader, Barnes, & Smeets, 1996, pp. 702-704; Leader, Barnes-Holmes, & Smeets, 2000; Leader & Barnes-Holmes, 2001, in press; Smeets, Leader, & Barnes, 1997).

Many RFT studies focus on non-equivalence relations, and it is here that a simple explanation in terms of stimulus correlations (involving either forward or backward associations) seems almost completely untenable. Consider, for example, a finding reported by Dymond and Barnes (1995). In this study, what we call a relational network was established using matching to sample procedures. The derived stimulus relations most relevant in the current context were as follows: C1 same as B1; B2 less than B1; C2 more than B1 (note the stimuli were nonsense syllables and thus were not related to each other along any consistent physical dimensions). A "one-response" discriminative function was then established for B1 using direct reinforcement. Subsequently, a "one-response" function emerged for C1, a "zero-response" function emerged for B2,

and a "two-response" function emerged for C2, without explicit training.

In effect, the directly reinforced "one-response" function established for B1 was transformed in accordance with the derived *same*, *less-than*, and *more-than* relations. B1 controlled one-response by direct training, and thus C1, which participated in a derived *same* relation with B1, also controlled one-response. B2, however, participated in a derived *less than* relation with B1, while C2 participated in a derived *more than* relation with B1. Thus, B2 controlled a zero-response, and C2 controlled two-responses.

We do not see how Tonneau's account can explain such data. During the training and testing that produced these performances each of the stimuli were correlated with each other (indirectly during training and directly during testing). If both B2 and C2 were correlated with B1, why would they acquire two different discriminative functions, each of which differ from the trained B1 function? Although difficult to explain using the concept of stimulus-correlations, data such as these are readily explained by RFT in terms of generalized operant response classes.

Other research findings that appear to stretch a stimulus correlational account are also handled with relative ease by RFT. Consider, for example, a study reported by Barnes, Hegarty, and Smeets (1997) in which matching to sample procedures were used to establish four, three-member equivalence classes (A1 B1 C1, A2 B2 C2, A3 B3 C3, A4 B4 C4). During one of the probe trials, the B1 and C1 stimuli were presented together as a compound sample stimulus, and two other compound stimuli were presented as comparisons; B2C2 and B3C4. Subjects reliably chose the B2C2 comparison over the B3C4 comparison, and the authors described this outcome as relating one equivalence relation to a second equivalence relation. This performance is readily predicted by RFT (see Stewart, Barnes-Holmes, Hayes, & Lipkens, 2001, for an extended discussion of relations among relations), but it is not obvious how the concept of stimulus correlations, per se, can accommodate these data. The individual stimuli in the B2C2 and B3C4 compounds were each equally correlated, during training and testing, with the two stimuli in the B1C1 compound (they were all incorrect stimuli in comparison sets), and thus there appears to be no basis on which to predict the choice of B2C2.

In another part of the Barnes et al. study (data not reported in the published version) the following trial-type was presented; B1C1 as sample with B2C2 and A1C4 as comparisons. In this case, the A1C4 choice contained a stimulus (A1) that had been *directly correlated*, during training, with both of the elements contained in the sample, whereas B2C2 contained two stimuli that were never correlated with either element. The subjects chose B2C2, as RFT would predict, because an equivalence relation is in a "frame of coordination" with an equivalence relation.

In case after case (we have presented only two, but the possible examples are many), a primary focus on stimulus correlations, per se, appears to run into trouble when we consider some of the data that has been gathered over the years under the rubric of RFT. We have made a similar argument against purely class-based accounts of equivalence (Barnes & Roche, 1996; Hayes & Barnes, 1997). It might be possible to save these more limited accounts by attempting to confine them only to certain kinds of studies of derived stimulus relations and not others, but this saves the accounts by killing them. It is not a theory of a narrow research preparation that we need. If RFT can explain equivalence and a host of other phenomena, while other accounts cannot even explain equivalence data such as those reported by Barnes, et al. (1997), parsimony alone suggests that RFT deserves the more serious consideration.

Conclusion

Equivalence research was rooted in the potential it provided for the analysis of language. The stultification of equivalence research comes from forgetting that mission. Tonneau appears to agree. However, like others before him (whether pro- or anti-equivalence), Tonneau fails to address the fundamental issues that seem to underlie language and cognition, providing instead an account that does not lead directly to powerful new approaches to human complexity.

We believe that it is time for behavioral theorists to face directly the challenge posed by other forms of derived stimulus relations, the transformation of functions through those relations, relational networks, relations among relations, and similar areas that RFT appears to handle with relative ease. We have shown that these phenomena lead directly to new empirical approaches to the analysis of such topics as metaphor, allegories, thinking, problem-solving, self, and similar topics (Hayes, Barnes-Holmes, & Roche, 2001). Furthermore, we have described how they lead directly to new approaches to psychological development, education, social behavior, psychopathology, psychotherapy, and even spirituality (Hayes, Barnes-Holmes, & Roche, 2001). Relational Frame Theory provides a relatively simple, operant account of the phenomenon of stimulus equivalence, but in addition it accounts for more complex phenomena as well, and in so doing gives behavior analysis a powerful new angle on human language and cognition itself.

To state it in simple terms, RFT claims that relating is an operant. This is an empirical matter, and must be so (see Barnes-Holmes, Barnes-Holmes, Roche, & Smeets, in press a, in press b; Healy, Barnes-Holmes, & Smeets, 2000). Behavior analysis is not a field worth having if it cannot determine, on empirical grounds, what is or is not an operant. Thus, the question "is relating an operant?" must be answerable in one of three ways: "no," "not enough data yet," or "yes." If the answer is "no" then one must say why, in clear conceptual or evidentiary terms. If the answer is "don't know yet" one must say what other evidence is needed. If the answer is "yes," then behavior analysis will be forever changed. If relating is an operant, it operates on operant and respondent processes themselves (Hayes, Fox, Gifford, Wilson, Barnes-Holmes, Healy, 2001; see also Barnes-Holmes & Barnes-Holmes, 2000; Barnes-Holmes, Barnes-Holmes, & Cullinan, 2000).

Evidence is mounting in favor of the RFT interpretation of equivalence responding, and relational responding more generally (see Hayes, Barnes-Holmes, & Roche, 2001, for a comprehensive review). The issue may not be closed as an empirical matter, but we believe that it is now clearly on the table. Of course, some may prefer to ignore these fundamental issues and hope instead that the corpse of stimulus equivalence will be reincarnated in a different guise (e.g., as the study of naming, stimulus control topographies, or stimulus correlations). We believe, however, that it is time to leave the graveside of stimulus equivalence, turn to the mission that this research area first provided, and attempt once more to scale the mountains of human language and cognition.

References

Barnes, D. (1994). Stimulus equivalence and relational frame theory. *The Psychological Record, 44*, 91-124.

Barnes, D., Hegarty, N., & Smeets, P. M. (1997). Relating equivalence relations to equivalence relations: A relational framing model of complex human functioning. *The Analysis of Verbal Behavior, 14*, 57-83.

Barnes, D., & Keenan, M. (1993). A transfer of functions through derived arbitrary and nonarbitrary stimulus relations. *Journal of the Experimental Analysis of Behavior, 59*, 61-81.

Barnes, D., & Roche, B. (1996). Relational frame theory and stimulus equivalence are fundamentally different: A reply to Saunders. *The Psychological Record, 46*, 489-508.

Barnes, D., Smeets, P.M., & Leader, G. (1996). Procedures for generating emergent matching performances: Implications for stimulus equivalence. In T.R. Zentall and P.M. Smeets (Eds.), *Stimulus class formation in humans and animals*(pp. 153-171). Holland: Elsevier.

Barnes-Holmes, D., & Barnes-Holmes, Y. (2000). Explaining complex behavior, Two perspectives on the concept of generalized operant classes. *The Psychological Record, 50*, 251-265.

Barnes-Holmes, D., Barnes-Holmes, Y., & Cullinan, V. (2000). Relational frame theory and Skinner's "Verbal Behavior:" A possible synthesis. *The Behavior Analyst, 23*, 69-84.

Barnes-Holmes, D., Keane, J., Barnes-Holmes, Y., & Smeets, P. M. (2000). A derived transfer of emotive functions as a means of establishing differential preferences for soft drinks. *The Psychological Record, 50*, 493-512.

Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (in press a). Exemplar training and a derived transformation of functions in accordance with symmetry. *The Psychological Record*.

Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (in press b). Exemplar training and a derived transformation of functions in accordance with symmetry: II. *The Psychological Record*.

Dymond, S., & Barnes, D. (1995). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness, more-than, and less-than. *Journal of the Experimental Analysis of Behavior, 64*, 163-184.

Dymond, S., & Barnes, D. (1996). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness and opposition. *The Psychological Record, 46*, 271-300.

Hall, G. (1996). Learning about associatively activated stimulus representations: Implications for acquired equivalence and perceptual learning. *Animal Learning and Behavior , 24*, 233-255.

Hayes, L. J. (1992). Equivalence as process. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 97-108). Reno, NV: Context Press.

Hayes, S. C., & Barnes, D. (1997). Analyzing derived stimulus relations requires more than the concept of stimulus class. *Journal of the Experimental Analysis of Behavior, 68*, 235-270.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). (Eds.), *Relational Frame Theory: A post-Skinnerian account of human language and cognition*. Plenum Press.

Hayes, S. C., Fox, E., Gifford, E. V., Wilson, K. G., Barnes-Holmes, D., & Healy, O. (2001). Derived relational responding as learned behavior. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational Frame Theory: A post-Skinnerian account of human language and cognition* (pp. 21-50). Plenum Press.

Hayes, S. C., Gifford, E. V., & Wilson, K. G. (1996). Stimulus classes and stimulus relations: Arbitrarily applicable relational responding as an operant. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 279-299). New York: Elsevier.

Healy, O., Barnes-Holmes, D., & Smeets, P. M. (2000). Derived relational responding as generalized operant behavior. *Journal of the Experimental Analysis of Behavior, 74*, 207-227.

Leader, G., Barnes, D., & Smeets, P. M. (1996). Establishing equivalence relations using a respondent-type training procedure. *The Psychological Record, 46*, 685-706.

Leader, G., Barnes-Holmes, D., & Smeets, P. M. (2000). Establishing equivalence relations using a respondent-type procedure III. *The Psychological Record, 50*, 63-78.

Leader, G., & Barnes-Holmes, D. (2001). Establishing fraction-decimal equivalence using a respondent-type training procedure. *The Psychological Record, 51*, 151-166.

Leader, G., & Barnes-Holmes, D. (in press). Matching-to-sample and respondent-type training as methods for producing equivalence relations: Isolating the critical variables. *The Psychological Record*.

Roche, B., & Barnes, D. (1996). Arbitrarily applicable relational responding and human sexual categorization: A critical test of the derived difference relation. *The Psychological Record, 46*, 451-475.

Roche, B., & Barnes, D. (1997). A transformation of respondently conditioned stimulus function in accordance with arbitrarily applicable relations. *Journal of the Experimental Analysis of Behavior, 67*, 275-300.

Roche, B., Barnes-Holmes, D., Smeets, P. M., Barnes-Holmes, Y., & McGeady, S. (2000). Contextual control over the derived transformation of discriminative and sexual arousal functions. *The Psychological Record,50*, 267-291.

Sidman, M. (1994). *Stimulus equivalence: A research story*. Boston: Authors Cooperative.

Skinner, B. F. (1938). *Behavior of organisms*. New York: Appelton-Century-Crofts.

Smeets, P.M., Leader, G., & Barnes, D. (1997). Establishing stimulus classes with adults and children using a respondent training procedure: A follow-up study. *The Psychological Record, 47*, 285-308.

Spetch, M . L., Wilkie, D. M., & Pinel, J. P. J. (1981). Backward conditioning: A reevalua-

tion of the empirical evidence. *Psychological Bulletin, 89*, 163-175.

Steele, D. L., & Hayes, S. C. (1991). Stimulus equivalence and arbitrarily applicable relational responding, *Journal of the Experimental Analysis of Behavior, 56*, 519-555.

Stewart, I., Barnes-Holmes, D., Hayes, S. C., & Lipkens, R. (2001). Relations among relations: Analogies, metaphors, and stories. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational Frame Theory: A post-Skinnerian account of human language and cognition* (pp. 73-86). Plenum Press.

Urcuioli, P. J. (1996). Acquired equivalences and mediated generalization in pigeon's matching-to-sample. In T.R. Zentall and P.M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 55-70). Holland: Elsevier.

Emergent matching to sample and equivalence relations

Harrie Boelens
Leiden University

Tonneau has presented a well-written, thoughtful, and erudite critique of the literature on stimulus equivalence. I will here react to two of the issues he has raised: the importance of studying emergent matching to sample, and the proper application of the mathematical "equivalence relation" concept.

The performances seen on tests of stimulus equivalence are no obvious conditioning effects. I propose that that is justification enough for studying them. Emergent matching to sample is a puzzle for conditioning theory. We can hope that working on this puzzle will lead to a revision of the theory, and that the revised theory will make sense of other emergent phenomena, such as those of generalized imitation, instruction following, and grammatical speech. For me, it is not necessary that the puzzle has a feature (such as "symbolic behavior") that makes it look relevant for something in everyday life. I agree with Tonneau that the cognitive journals present many facts that do not have obvious behavior-analytic interpretations, but does it matter? Studying them will be worthwhile too! I find it a pity only that there is not much diversity in this work. Working on many different puzzles at the same time might be more fruitful than the massive concentration on emergent matching to sample that we see nowadays.

The applications of the mathematical "equivalence relation" concept to emergent matching to sample often show misunderstandings. Tonneau will agree with me. I will here develop his point further, in the hope of producing more clarity. Following Tonneau, define *xRy* to mean that comparison *y* is reliably selected in the presence of sample *x*. We could suggest that *R* is an equivalence relation (i.e., a relation that is reflexive, symmetric, and transitive). This suggestion has many consequences. I will here consider five that are not often appreciated. First, it does not make sense to call the individual relations that are trained or tested equivalence relations. Suppose we have taught a subject conditional relations A1B1, A2B2, B1C1, and B2C2. We can then ask if *the* conditional relation is an equivalence relation (for this subject, at this time). It does not make sense to ask if A1B1, A2B2, etc. are equivalence relations. Second, the relation *R* applies to matching to sample, not to conditional discrimination in general. Suppose A1 is a red light, A2 a green light, B1 a low tone, and B2 a high tone. A rat's lever

Address correspondence concerning this article to E-mail: boelens@fsw.leidenuniv.nl