The Psychological Record, 2002, 52, 305-314

STUDYING STIMULUS EQUIVALENCE: DEFENSE OF THE TWO-CHOICE PROCEDURE

HARRIE BOELENS
Leiden University

Sidman (1987) and Carrigan and Sidman (1992) have advised against the use of two-choice procedures in studies of emergent matching to sample. They argue that in two-choice, as opposed to multiple-choice, procedures: (1) It is more difficult to make sure that the baseline conditional relations have been established; (2) There is a greater chance that test outcomes are not related to the baseline conditional relations; and (3) The predictions of stimulus equivalence are less clear. In response to the first two arguments, I argue that they refer to technical difficulties that can easily be handled within the two-choice procedure itself. In response to the third argument, I argue that the formulation of Carrigan and Sidman is a new account, that it is inconsistent with the old account of Sidman and Tailby (1982), and that the two-choice procedure goes well with the old account. Further, I argue that there are no strong reasons for adopting the new account, and that the recommendation of using three choices in particular is problematic.

Nowadays, many behavior analysts study transfer phenomena in human matching to sample. An important concept in this research is the *conditional relation* (Sidman & Tailby, 1982). Suppose that sample A1 alternates irregularly with other samples, and that the same set of comparisons, including B1, is arranged in the presence of each sample. If the choice of B1 predominates in the presence of A1, and not in the presence of at least one of the other samples, then conditional relation A1B1 is said to be shown.

Much of this research examines the effects of training a set of interrelated conditional relations, such as {AB, AC} or {AB, BC, CD}. It has been found that the training of such a set can lead to a variety of other, untrained conditional relations. A rough characterization of the findings is the statement that the conditional relation is an *equivalence relation*, that is, a relation that is reflexive, symmetric, and transitive (Saunders &

I thank Paul Smeets for his comments on previous versions of this manuscript. Address correspondence to Harrie Boelens, Department of Psychology, Leiden University,

Green, 1992; Sidman & Tailby, 1982). When the conditional relation is an equivalence relation, stimuli form classes, such that conditional relations are shown between stimuli in the same class. Stimuli in the same class are called *equivalent*, and the actual existence of classes of equivalent stimuli is called *stimulus equivalence*.

The emergence of untrained conditional relations has been studied with a wide variety of procedures. In this paper, the expression *two-choice procedure* will stand for a matching-to-sample procedure that arranges two comparisons in the presence of each sample; the expression *multiple-choice procedure* for a procedure that arranges more than two comparisons. Evidence for stimulus equivalence has been obtained with both two-choice (e.g., Lynch & Green, 1991; Pilgrim & Galizio, 1990) and multiple-choice procedures (e.g., Sidman & Cresson, 1973; Sidman & Tailby, 1982). Until 1987 these procedures were valued in the same way. Since then, however, Sidman (1987) and Carrigan and Sidman (1992) have advised against the use of two-choice procedures (see also Sidman, 1994, 2000). They give three arguments for this. I will respond to these arguments and attempt to show that none points to an important weakness of the two-choice procedure. Sidman (1987, pp. 14-15) saw the possibility of responses to the first two arguments. Nevertheless, he wished to maintain his conclusion that multiple-choice procedures should be preferred. Therefore, the third argument seems especially important.

Arguments Against Two-Choice Procedures

First, Sidman (1987, pp. 14-15) argues that in two-choice, as opposed to multiple-choice, procedures, it is more difficult to make sure that the baseline conditional relations have been established. This argument is concerned with how to demonstrate a conditional relation. A conditional relation involves the reliable choice of a particular comparison in the presence of a particular sample. This requires the use of stringent criteria, and these should be applied to all individual conditional relations. Requiring 75% correct across conditional relations A1B1 and A2B2 in a two-choice task is not enough. Researchers are well aware of this. In two-choice procedures, they routinely require high percentages of correct responding in the presence of each sample. For example, Lynch and Green (1991) presented samples A1 and A2 each eight times in blocks of 16 trials. Conditional relations A1B1 and A2B2 were assumed when B1 was chosen in the presence of A1, or B2 in the presence of A2, on at least 31 trials, in two consecutive blocks. Clearly, making sure that the baseline relations have been established is not really a problem, even in two-choice procedures. It is only a technical difficulty that can be, and usually is, solved within the two-choice procedure itself.

Second, Sidman (1987, p. 15) argues that in two-choice, as opposed to multiple-choice procedures, there is a greater chance that positive tests of stimulus equivalence are not related to the baseline conditional relations. This point is more subtle, and the problem may be

underestimated. Suppose relations A1B1 and A2B2 are established in a two-choice procedure. Thereafter, relations B1A1 and B2A2 are tested and indeed found. One might now want to believe that B1A1 and B2A2 are shown because of the training that produced A1B1 and A2B2, but this need not be the case. There need be no causal relationship here.

Again, researchers are aware of the problem, and they deal with it in two ways. The usual solution is to replicate the same kind of test with other stimuli. This is often done with interrelated stimulus sets. For example, AB and AC relations may be trained and BA and CA relations tested, which yields two symmetry tests (Pilgrim & Galizio, 1990). Sometimes independent conditional relations are used (e.g., AB training followed by BA tests, CD training followed by DC tests, etc.; see Boelens, Van den Broek, & Van Klarenbosch, 2000; Stromer & Osborne, 1982). The second solution is to counterbalance the pairings of stimuli in trained conditional relations. For example, relations A1B1 and A2B2 may be trained in one group of subjects, and A2B1 and A1B2 in another (Hogan & Zentall, 1977; Stromer & Osborne, 1982). Relations B1A1 and B2A2 might then be shown on symmetry tests in the first group, and B1A2 and B2A1 in the second. This is proof of dependence of the tested on the trained conditional relations. A variation on this strategy is to randomly assign stimuli to stimulus designations (Markham & Dougher, 1993). This has roughly the same effect as counterbalancing. To continue the example given above: one subject might receive training in A1B1 and A2B2, another in B1A2 and B2A1, still another in A1B2 and A2B1, etc., and in all these cases the symmetric versions of the trained relations are tested. When symmetry is shown by all or most subjects, the relations shown vary with what has been taught, just as in the counterbalancing procedure. Clearly, then, the problem of demonstrating the dependence of test on training results in two-choice procedures can be tackled in various ways, and researchers do apply these tactics.

Third, Sidman (1987) and Carrigan and Sidman (1992) argue that the predictions of stimulus equivalence are less clear for two-choice than for multiple-choice procedures. This argument refers to selection and rejection, two kinds of stimulus control that are possible in matching to sample. This will be discussed separately below.

Selection and Rejection in Matching to Sample

Sidman (1987) and Carrigan and Sidman (1992) distinguish two kinds of stimulus control in matching to sample (see also Dube & McIlvane, 1996; Stromer & Osborne, 1982). In *selection* or sample/S+ control, responding is controlled by the sample and the correct comparison; the incorrect comparisons do not exert control. In *rejection* or sample/S- control, responding is controlled by the sample and an incorrect comparison; the correct comparison does not exert control. Selection can be viewed as a relation between a sample and a correct comparison, and rejection as a relation between a sample and an incorrect comparison. Carrigan and Sidman (1992) use the terms

controlling relation and *conditional relation* interchangeably for these two relations, but it may be better to use the term controlling relation only. This will avoid confusion with the conditional relation as it was defined by Sidman and Tailby (1982).

Selection and rejection could both be equivalence relations. For example, suppose that rejection is transitive, that comparison B2 is rejected in the presence of sample A1, and that comparison C1 is rejected in the presence of sample B2. Comparison C1 will then be rejected in the presence of sample A1. Carrigan and Sidman (1992) derive equivalence predictions for selection and rejection on a large variety of two-choice tests. The predictions turn out to be the same on symmetry tests and on some of the more complex tests, but different on reflexivity and transitivity tests, and on other complex tests. Therefore, there are no predictions for the second kind of test unless the controlling relation (selection or rejection) is known in advance.

Carrigan and Sidman (1992) view selection and rejection as "means of performing a particular conditional discrimination" (p. 188), and they are willing to make assumptions about the occurrence of selection or rejection on the basis of their efficiency. In two-choice procedures, selection is just as efficient as rejection, because selection of the correct comparison and rejection of the incorrect comparison both lead to correct responding. Therefore, Carrigan and Sidman suggest that one should reckon with both, which prevents predictions of stimulus equivalence (on tests of reflexivity, transitivity, etc.). In multiple-choice procedures, selection is the more efficient procedure. Here, a single sample/S+ relation still works, but now there are at least two incorrect stimuli, and therefore at least two sample/S- relations are needed. Carrigan and Sidman therefore suggest that selection predominates in multiple-choice procedures, which leads to clear predictions for these procedures. This makes Carrigan and Sidman favor multiple-choice procedures. These procedures can be inconvenient, however, and the more so the larger the number of choices. As a compromise, Carrigan and Sidman recommend using three-choice procedures.

In response to this, I will argue that the formulation of Carrigan and Sidman is a new account, that it is inconsistent with the old account of Sidman and Tailby (1982), and that the two-choice procedure goes well with the old account. Further, I will argue that there are no strong reasons for adopting the new account, and that the recommendation of using three choices in particular is problematic.

Response to the Third Argument

The formulation of Carrigan and Sidman (1992) deviates in important ways from the account originally proposed by Sidman and Tailby (1982) and developed further by Sidman (1986, 1990, 1992, 2000). The deviations stem from the new conditional relation concept. In the formulation of Sidman and Tailby (1982, see above), a conditional relation is shown when the choice of a particular comparison predominates in the presence of a particular sample.

This is "directly observable by reference to the subject's ongoing interactions with the procedure. Testing for the existence of a conditional relation requires no modification of the establishing procedure" (p. 5). Sidman and Tailby suggest that the conditional relation can be an equivalence relation, and they show that this leads to clear predictions. These are the same for two- and multiple-choice procedures.

The new account applies the equivalence concept to a new definition of the conditional relation. The conditional relation has now become a controlling relation, selection or rejection:

> The conditional discrimination procedure may generate conditional relations between samples and either positive or negative comparisons. We shall call the relation between samples and positive comparisons the *select* or *Type S* relation ... Alternatively, the same performance could reflect a relation between samples and incorrect comparisons. We shall call this the *reject* or *Type R* relation. (p. 185)

The proof of selection or rejection requires tests because the relevance or irrelevance of certain comparison stimuli needs to be demonstrated. According to the new account, selection and rejection can be equivalence relations (p. 185). This account does not make predictions for some standard tests of stimulus equivalence, unless the controlling relations are known. When only rejection is present, the predictions for these tests are different from those of Sidman and Tailby (1982). In other words, the two accounts are inconsistent.

The inconsistency of the two accounts is an important point, but it may be overlooked. The old and the new account both apply the mathematical equivalence concept to a behavioral conditional relation concept. It might therefore be thought that the new account is only a refinement of the old one. Instead, it should be acknowledged that the two are inconsistent and that there are no problems with the two-choice procedure in the old account.

Is it good research strategy to change to the new account, with its recommendation of arranging more than two comparisons? In favor of this, one might suggest that the equivalence of rejection explains some of the exceptions to the old account. It has sometimes been found that emergent conditional relations do not reverse when some of the baseline relations have been reversed (Pilgrim & Galizio, 1990, 1995; Saunders, Saunders, Kirby, & Spradlin, 1988). To explain this, Carrigan and Sidman have suggested that the nature of the controlling relation might change during the reversal of baseline relations. Originally, correct comparisons might be selected. Later, some of these become incorrect in the presence of the same samples, and these comparisons might now be rejected. Johnson and Sidman (1993) have obtained support for this explanation. They attempted to deliberately produce rejection, and obtained test results that were completely in agreement with the idea that rejection can be an equivalence relation. Many of these results were inconsistent with Sidman and Tailby (1982).

These are important findings, but we can give them a place without adopting the new account. Instead, we could accept the findings as indications of limits on stimulus equivalence in the old sense. That is, our strategy could be to maintain the old conditional relation concept, and to determine the conditions for stimulus equivalence in the old sense. This seems to be better. It is simple, and it means we can continue to talk about emergent matching to sample as we did. Further, the occurrence of rejection is only one account of the failures to reverse (for others, see Garotti, de Souza, de Rose, Molina, & Gil, 2000; Pilgrim & Galizio, 1996; Saunders, Drake, & Spradlin, 1999) and other exceptions to stimulus equivalence are not so readily accounted for. These include findings with animals and human children on two-choice symmetry tests (e.g., Boelens & Van den Broek, 2000; Dugdale & Lowe, 2000; Lipkens, Kop, & Matthijs, 1988), and with human children and adults on multiple-choice tests (e.g., Healy, Barnes-Holmes, & Smeets, 2000, Experiment 2; Lazar, Davis-Lang, & Sanchez, 1984). In light of this, changing to the new account seems premature.

In addition, there are problems with the new account, and with the recommendation of arranging more than two comparisons. Problems with the new account follow from the fact that it refers to kinds of stimulus control that are not directly observable in the matching-to-sample performances. One consequence of this is that old data can be accounted for only with additional assumptions. For example, evidence for stimulus equivalence has often been found in two-choice procedures. To explain this with the new account, we will have to assume either that rejection did not occur (even though just as efficient as selection), or that rejection was not an equivalence relation in these studies. Another consequence is that new data will provide incomplete evidence unless stimulus control tests give appropriate results. Otherwise, it can always be argued that exceptions to stimulus equivalence (in the old sense) occurred for some reason other than rejection.

Carrigan and Sidman suggest that the presence of rejection can be concluded from the outcome of standard equivalence tests, but this is not satisfactory. When we work this way we will be confident about rejection only when many different tests agree (as in Johnson & Sidman, 1993). The rejection account becomes much weaker when some, but not all, tests agree (as in Pilgrim & Galizio, 1990). There is a circularity here: Rejection is inferred from standard equivalence tests, and at the same time it is used to explain the outcomes of those tests. To break this circularity decisively, one needs to do tests of stimulus control.

Another problem with the new account is that it may promote a cognitive or mediational approach. Carrigan and Sidman (1992) proposed to use the words selection and rejection "merely as shorthand labels to indicate which comparison stimulus, along with the sample, controls the subject's performance" (p. 185). It is not easy, however, to maintain these definitions. The two words can easily come to stand for activities that precede or accompany choices, and many statements of

Carrigan and Sidman may promote this. When they explain selection they speak of *selecting and touching* a comparison (see p. 187ff; my italics), and when they discuss findings that agree with selection they state that "the subject might always have chosen correctly without ever having identified the negative stimulus on any baseline or test trial" (p. 186). When they explain rejection, they often speak of *rejecting* one comparison *and touching* another (see p. 187ff; my italics), and when they discuss findings that agree with rejection they state that "the subject need never have identified the stimuli that were actually touched" (p. 186). Further, when explaining what might happen on a test of stimulus control, they state that a subject "might recall that whenever comparison B2 had been there to be rejected, B1 had been available to be touched" (p. 186). These statements can make one look for activities other than choosing comparisons, such as observing behavior or speech. For example, one might want to conclude that rejection of a comparison has taken place if the comparison was scanned, but subsequently not chosen. This would be beside the point when the original stimulus-control definitions are maintained.

Two problems with the recommendation of arranging more than two comparisons are the following. First, the recommendation of using three choices is weak. It is based on the assumption that rejection is not important in three-choice procedures. This requires proof, and the evidence that we have until now points in the opposite direction. A number of investigators have found that subjects can be taught to respond away from incorrect stimuli in three-choice procedures (Innis, Lane, Miller, & Critchfield, 1998; Serna, Wilkinson, & McIlvane, 1998; Wilkinson & McIlvane, 1997). This has been found in studies with a blank or none option. In these studies, the blank or "none" option was introduced after subjects had learned a standard three-choice arbitrary matching task. On some trials, the new option was correct (substituted for the correct comparison); on others, it was incorrect (substituted for an incorrect comparison). Correct responding was easily maintained in these procedures, and this was true for college students (Innis et al., 1998), 3- to 5-year-old children (Wilkinson & McIlvane, 1997), and severely retarded children and adolescents (Serna et al., 1998). The choice of the new option, when it occurs beside incorrect comparisons, demonstrates control by incorrect stimuli. This finding suggests that rejection is present in conventional three-choice arbitrary matching as well. It does not prove it, because it was deliberately taught here, with a special comparison stimulus.

The second problem is that by using multiple-choice procedures we will fail to notice exceptions to stimulus equivalence (in the old sense) that might occur in two-choice procedures for reasons other than rejection. One of these reasons is connected with the occurrence of consistent responding (for discussion, see Devany, Hayes, & Nelson, 1986; Sidman, 1992). There are many ways that subjects can show consistency on standard tests of stimulus equivalence, but for the sake of simplicity, assume that subjects either choose the comparison dictated by stimulus equivalence (as defined by Sidman and

Tailby, 1982), or choose away from that comparison. In two-choice procedures, both of these "rules" lead to consistency; in three-choice procedures, only stimulus equivalence is available. Thus, two-choice procedures might lead to exceptions to stimulus equivalence, because there are ways other than stimulus equivalence to be consistent in that situation. If that is the case then consistency is more basic or primitive than stimulus equivalence. Stimulus equivalence is then only a way to be consistent that will be facilitated when there are no alternatives available. We will never find this out if we restrict the study of stimulus equivalence to multiple-choice procedures.

Conclusion

Sidman (1987) and Carrigan and Sidman (1992) have given three arguments against the use of two-choice procedures in studies of emergent matching to sample. In response to the first two arguments, I have argued that they refer to technical difficulties that can easily be handled within the two-choice procedure itself. In response to the third argument, I have argued that the formulation of Carrigan and Sidman is a new account, that it is inconsistent with the old account of Sidman and Tailby (1982), and that the two-choice procedure goes well with the old account. Further, I have argued that there are no strong reasons for adopting the new account, and that the recommendation of using three choices in particular is problematic.

Sidman's latest formulation (Sidman, 2000) defines the behavior that is to be accounted for in the same way as Sidman and Tailby (1982). The behavior is called "conditional relation," "conditional discrimination," "stimulus pair," or "event pair," and these expressions refer either to the choice of a particular comparison in the presence of a particular sample, or to a particular response to a particular stimulus. The equivalence concept is applied to this behavior, rather than to selection or rejection. It would be a pity if the testing of this formulation would be restricted by the recommendation of arranging more than two comparisons.

References

BOELENS, H., & VAN DEN BROEK, M. (2000). Influencing children's symmetric responding in matching-to-sample tasks. *The Psychological Record*, 50, 655-669.

BOELENS, H., VAN DEN BROEK, M., & VAN KLARENBOSCH, T. (2000). Symmetric matching to sample in 2-year-old children. *The Psychological Record*, 50, 293-304.

CARRIGAN, P. F., JR., & SIDMAN, M. (1992). Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the Experimental Analysis of Behavior*, 58, 183-204.

DEVANY, J. M., HAYES, S. C., & NELSON, R. O. (1986). Equivalence-class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior*, 46, 243-257.

DUBE, W. V., & MCILVANE, W. J. (1996). Some implications of a stimulus control topography analysis for emergent behavior and stimulus classes. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 197-218). Amsterdam: Elsevier.

DUGDALE, N., & LOWE, C. F. (2000). Testing for symmetry in the conditional discriminations of language-trained chimpanzees. *Journal of the Experimental Analysis of Behavior*, 73, 5-22.

GAROTTI, M., DE SOUZA, D. G., DE ROSE, J. C., MOLINA, R. C., & GIL, M. S. A. (2000). Reorganization of equivalence classes after reversal of baseline relations. *The Psychological Record*, 50, 35-48.

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.

HOGAN, D. E., & ZENTALL, T. R. (1977). Backward associations in the pigeon. *American Journal of Psychology*, 90, 3-15.

INNIS, A., LANE, S. D., MILLER, E. R., & CRITCHFIELD, T. S. (1998). Stimulus equivalence: Effects of a default-response option on emergence of untrained stimulus relations. *Journal of the Experimental Analysis of Behavior*, 70, 87-102.

JOHNSON, C., & SIDMAN, M. (1993). Conditional discrimination and equivalence relations: Control by negative stimuli. *Journal of the Experimental Analysis of Behavior*, 59, 333-347.

LAZAR, R. M., DAVIS-LANG, D., & SANCHEZ, L. (1984). The formation of visual stimulus equivalences in children. *Journal of the Experimental Analysis of Behavior*, 41, 251-266.

LIPKENS, R., KOP, P. F. M., & MATTHIJS, W. (1988). A test of symmetry and transitivity in the conditional discrimination performances of pigeons. *Journal of the Experimental Analysis of Behavior*, 49, 395-409.

LYNCH, D. C., & GREEN, G. (1991). Development and crossmodal transfer of contextual control of emergent stimulus relations. *Journal of the Experimental Analysis of Behavior*, 56, 139-154.

MARKHAM, M. R., & DOUGHER, M. J. (1993). Compound stimuli in emergent stimulus relations: Extending the scope of stimulus equivalence. *Journal of the Experimental Analysis of Behavior*, 60, 529-542.

PILGRIM, C., & GALIZIO, M. (1990). Relations between baseline contingencies and equivalence probe performances. *Journal of the Experimental Analysis of Behavior*, 54, 213-224.

PILGRIM, C., & GALIZIO, M. (1995). Reversal of baseline relations and stimulus equivalence: I. Adults. *Journal of the Experimental Analysis of Behavior*, 63, 225-238.

PILGRIM, C., & GALIZIO, M. (1996). Stimulus equivalence: A class of correlations, or a correlation of classes? In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 173-195). Amsterdam: Elsevier.

SAUNDERS, R. R., DRAKE, K. M., & SPRADLIN, J. E. (1999). Equivalence class establishment, expansion, and modification in preschool children. *Journal of the Experimental Analysis of Behavior*, 71, 195-214.

SAUNDERS, R. R., & GREEN, G. (1992). The nonequivalence of behavioral and mathematical equivalence. *Journal of the Experimental Analysis of Behavior*, 57, 227-241.

SAUNDERS, R. R., SAUNDERS, K. J., KIRBY, K. C., & SPRADLIN, J. E. (1988). The merger and development of equivalence classes by unreinforced conditional selection of comparison stimuli. *Journal of the Experimental Analysis of Behavior*, 50, 145-162.

SERNA, R. W., WILKINSON, K. M., & MCILVANE, W. J. (1998). Blank-comparison assessment of stimulus-stimulus relations in individuals with mental retardation: A methodological note. *American Journal on Mental Retardation*, 103, 60-74.

SIDMAN, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213-245). Hillsdale, NJ: Erlbaum.

SIDMAN, M. (1987). Two choices are not enough. *Behavior Analysis*, 22, 11-18.

SIDMAN, M. (1990). Equivalence relations: Where do they come from? In D. E. Blackman & H. Lejeune (Eds.), *Behaviour analysis in theory and practice: Contributions and controversies* (pp. 93-114). Hove: Lawrence Erlbaum Associates.

SIDMAN, M. (1992). Equivalence relations: Some basic considerations. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 15-27). Reno, NV: Context Press.

SIDMAN, M. (1994). *Equivalence relations and behavior: A research story.* Boston: Authors Cooperative.

SIDMAN, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, 74, 127-146.

SIDMAN, M., & CRESSON, O., JR. (1973). Reading and crossmodal transfer of stimulus equivalences in severe retardation. *American Journal of Mental Deficiency*, 77, 515-523.

SIDMAN, M., & TAILBY, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, 37, 5-22.

STROMER, R., & OSBORNE, J. G. (1982). Control of adolescents' arbitrary matching-to-sample by positive and negative stimulus relations. *Journal of the Experimental Analysis of Behavior*, 37, 329-348.

WILKINSON, K. M., & MCILVANE, W. J. (1997). Blank comparison analysis of emergent symbolic mapping by young children. *Journal of Experimental Child Psychology*, 67, 115-130.