Paper:
Reed, P. & Quigley, M. (2019). Over-selectivity decreases with increased training: A role for within-compound associations. *Acta Psychologica, 198*, 102868
http://dx.doi.org/10.1016/j.actpsy.2019.102868

Over-selectivity decreases with increased training: A role for within-compound associations

Phil Reed and Martyn Quigley

Swansea University

Correspondence address: Phil Reed,

Department of Psychology,

Swansea University,

Singleton Park,

Swansea, SA2 8PP, U.K.

Short title: Overshadowing and training.

Cite as: Reed, P., & Quigley, M. (2019). Over-selectivity decreases with increased training: A role for within compound associations. *Acta Psychologica*, **198**,
doi: 10.1016/j.actpsy.2019.102868

Abstract

Over-selectivity occurs when one element of a complex-stimulus controls behavior at the expense of other equally elements of that stimulus; a phenomenon common in populations subject to cognitive challenge. However, lack of theoretically-based analysis, may have hindered understanding and remediation of the practically-important over-selectivity phenomena. Current studies examined whether associative theories applied to overshadowing, a similar phenomenon in the context of conditioning experiments, could be applied to over-selectivity effects to open theoretical analysis of over-selectivity. Three experiments investigated whether length of training impacts over-selectivity in the same way as overshadowing, which has theoretical implications for understanding that latter phenomenon. All studies employed variants of a judgment procedure in which participants had to judge the relationship between a predictor and an outcome, and the predictors were presented either on their own, or in compound with another predictor. In all studies, the elemental cue (A) was rated similarly to one of the components of the compound (B), but higher than the other component (C). The difference in the extent to which the components of the compound (B and C) were judged as predictors became smaller as levels of training increased, which is an effect that is also seen in discrimination learning studies of over-selectivity. Moreover, it was apparent that as the strength of the within-compound association increased, the level of over-selectivity decreased. These results are similar to those seen for overshadowing, and are discussed with respect to the possible associative mechanisms controlling over-selectivity.

Keywords: overshadowing; length of training; judgment; within-compound associations.

One form of cue interaction effect is known as stimulus over-selectivity (Lovaas & Schreibman, 1971; Lovaas, Schreibman, Koegel, & Rehm, 1971; Reed & Gibson, 2005; Reed, Reynolds, & Fermandel, 2012), and refers to situation where behavior is controlled by only one out of a number of equally-important elements of the environment (e.g., Dube & McIlvane, 1999; Lovaas, Koegel, & Schreibman, 1979; Reed & Gibson, 2005; see Dube, 2009 for a review). The over-selectivity effect is widespread across many disabilities, including Autism Spectrum Disorders (ASD; Allen & Fuqua, 1985; Hedbring & Newsom, 1985; Huguenin, 1997; Lovaas & Schreibman, 1971; Reed, Broomfield, McHugh, McCausland, & Leader, 2009), learning disabilities (Bailey, 1981; Dube & McIlvane, 1999, Gersten, 1983; Lovaas et al., 1971; Schneider & Salzberg, 1982; Stromer, McIlvane, Dube, & Mackay, 1993), and acquired neurological damage (Wayland & Taplin, 1982, 1985), as well as being observed in the elderly (Kelly, Leader, & Reed, 2016; McHugh & Reed, 2007). However, over-selectivity also can occur in typical/healthy adults, especially under conditions of high cognitive demands or concurrent activity (e.g., Broomfield, McHugh, & Reed, 2008a, 2008b; 2010; Dube, Balsamo, Fowler, Dickson, Lombard, & Tomanari, 2006; Reed, 2006; Reed & Gibson, 2005; Reynolds & Reed, 2011a, 2011b, 2012; Reynolds, Watts, & Reed, 2012).

In practical terms, over-selectivity impairs individual’s ability to simultaneously monitor several cues, and can result in a variety of problems, such as: social skills deficits (e.g., Schreibman & Lovaas, 1973), language, communication, and speech problems (e.g., Chiang & Carter, 2008; Koegel, Schreibman, Britten, & Latinen, 1979; Lovaas et al., 1979; Schreibman et al., 1986), and emotional and behavioral problems (e.g., Cook, Anderson, & Rincover, 1982; Lovaas et al., 1979; Varni et al., 1979). However, although the effect is widespread, and has important practical implications, very little is known regarding the theoretical underpinnings of this ubiquitous effect.

To the degree that over-selectivity is discussed in theoretical terms, it is often attributed to differentially-disrupted attentional processing in some populations (Dube, 1999; Reed, Hawthorn, Bolger, Meredith, & Bishop, 2012). These views suggest that individuals who display over-selectivity cannot attend to all of the stimulus elements available, and consequently learn only about a subset of these elements (Dube & McIlvane, 1999). Over-selectivity, typically, is not discussed in terms of established learning models that may have greater generality in explaining the effect itself. A lack of more theoretically-based analysis of over-selectivity may have precluded application of a range of learning theories to a very wide range of practically important phenomena (see Lovaas et al., 1971; Reed, 2011).

One such possibility concerns exploring over-selectivity effects in the light of better understood cue-competition effects, such as stimulus overshadowing (e.g., Mackintosh, 1976); that is, attenuation of responding to one conditioned stimulus as a result of the presence of a second, often more powerful or salient, conditioned stimulus, in a compound conditioning design (AB+). In extending theoretical analyses developed from overshadowing procedures to over-selectivity procedures, further potential cognitive deficits implicated in producing the latter effect might be highlighted.

It is important to note that, although over-selectivity seems *prima face* to share similarities with overshadowing, there are operational distinctions between the two that may negate assumptions that similarities exist at the level of mechanism. Procedurally, overshadowing is often revealed when compound conditioning results in less responding for an element when presented alone following compound training, compared to when they were trained as individual elements (i.e., less responding to X after AX+ than to X after X+). However, over-selectivity occurs when there is more responding to A after AB+/CD- training, compared to B after AB+/CD-, even when A and B are matched for salience. Thus, over-selectivity is defined by the relative relationship of the control acquired by elements A and B after simultaneous discrimination training, whereas overshadowing is defined by reference to a reduced amount of control exerted by one stimulus (X) from the compound (AX+), relative to the control exerted when that element is conditioned individually (X+).

Given this, while the adoption of theoretical perspectives developed for overshadowing may help, it is unclear whether they would automatically apply to over-selectivity procedures. Moreover, Maes et al. (2016) have questioned the generality of cue competition phenomena by documenting numerous experiments that failed to find evidence for a blocking effect (i.e., impaired learning of a CS-US association if the CS is presented simultaneously during conditioning with a different CS that has already been associated with the US; Kamin, 1969). Thus, it cannot be assumed that such effects will be seen in every procedure, and that theoretical analyses developed for one area will automatically be applicable to another. As there are questions regarding the generality of such effects, exploring procedural variables and boundary conditions under which these cue competition effects are observed, and their potential mediating mechanisms, is fundamental prior to assuming that the two are the same. The current series of experiments aimed to explore one of these empirical similarities – the impact of extended training – to bring together the sorts of procedures and explanations offered in terms of studying over-selectivity and overshadowing.

A number of experiments have noted that overshadowing and over-selectivity both decrease with extended training (e.g., Bellingham & Gillette, 1981; Reynolds & Reed, 2011a; Stout, Arcediano, Escobar, & Miller, 2003). For example, Reynolds and Reed (2011a) noted that over-selectivity is greater with fewer rather than greater numbers of training trials. Similarly, Stout et al. (2003) noted that a greater number of compound conditioning trials diminished an overshadowing effect in rats, compared to fewer compound conditioning trials. Similar effects have been noted with blocking procedures (see Azorlosa & Cicala, 1988, James & Wagner, 1980; Kamin, 1969). The relationship between length of training and the size of the cue competition effect is important to establish in order to understanding over-selectivity effects within an associative learning framework, as evidence that other cue competition effects reduce as a function of training has been the subject of some theoretical debate (e.g., Denniston, Savastano, & Miller, 2001; Wheeler & Miller, 2008). However, prior to further discussion of this theoretical point in relation to over-selectivity, the reduction of the over-selectivity effect needs to be reliably established across procedures employed for humans in the study of cue competition that are more akin to those used within an associative learning framework. This would address problems in combining results from nonhuman overshadowing and human over-selectivity experiments, which use different procedures from one another, making the theoretical implications of any parallels unclear. Previously, Quigley and Reed (2017) have studied over-selectivity in a human judgment procedure, that is similar to procedures used to study human cue-competition (e.g., Van Hamme & Wasserman, 1994). Although this human judgment over-selectivity procedure is not identical to those used in previous cue competition studies, it bears strong similarities to the type of tasks these latter procedures employ, and, given it has demonstrated an over-selectivity effect, it was deemed suitable to employ for the current purposes.

Typically, stimulus over-selectivity is studied using a simultaneous discrimination procedure involving two compound stimuli: following an initial acquisition phase (AB+ CD-), the elements of the previously reinforced compound (i.e., AB) are subsequently tested separately in extinction, and one of the elements (e.g., A) is found to control behavior to a much greater extent than the other previously reinforced element (see Dube, Lombard, Farren, Flusser, Balsamo, & Fowler, 1999; Koegel & Scheribman, 1977; Lovaas, Berberich, Perloff, & Schaeffer, 1966; Lovaas et al., 1971; Lovaas & Schriebman, 1971; Reed & Gibson, 2005). The current series of studies explored whether a stimulus over-selectivity effect (e.g., Lovaas et al., 1971; Reed et al., 2009; Wayland & Taplin, 1985) would occur in a judgment procedure that is more commonly employed in the study of human cue competition effects, such as a judgment tasks (e.g., Price & Yates, 1993; Van Hamme & Wasserman, 1994). The use of such a procedure would also allow exploration of whether any stimulus over-selectivity effect obtained using this more traditional judgment paradigm wanes with extended training, as it appears to do in studies of over-selectivity (Reynolds & Reed, 2011a; 2011b), which has not been established for this population.

Experiment 1

The first experiment sought to investigate whether any over-selectivity effect (i.e., the tendency of one stimulus to control behavior at the expense of the other) would become smaller as training progressed using a variation of the judgment procedure employed by Van Hamme & Wasserman (1994). In the current version of the procedure, participants play the role of a health professional asked to rate the likelihood that various symptoms predict the presence of a particular illness. Each participant was exposed to a number of 'patients' (trials) who display various symptoms. During training, participants were asked to learn which symptoms predict the presence of an illness. The symptoms associated with an illness could either be presented elementally (A+), where only one symptom predicts an illness; or in compound (BC+), where two symptoms together always predicted an illness. If over-selectivity occurred (Lovaas et al., 1971; Reed & Gibson, 2005), then one of the stimuli (symptoms) in the compound would come to elicit a positive response regarding the presence of the illness to a much greater extent than the other, despite their equal validity in predicting the illness. If the effect reduces over time (Azorlosa & Cicala, 1988; Reynolds & Reed, 2011b; Stout et al., 2003), then the difference in the extent to which the elements of the compound are judged to predict the illness should decrease as training continues. This would serve both to demonstrate an over-selectivity effect in a judgment paradigm of the used to explore human associative learning, and show that the over-selectivity effect reduces with training, which has been noted for rats in overshadowing procedures (Bellingham & Gillette, 1981; Stout et al., 2003), and in human simultaneous discrimination over-selectivity procedure (Reynolds & Reed, 2011a), but not in human judgment procedures.

Method

Participants

Fifteen participants (5 male and 10 female) aged between 19 and 25 years took part. None of the participants had any prior experience with the task, and all were volunteers were recruited in the Department of Psychology. Ethical approval for the study was obtained through the Department of Psychology Ethics Committee.

Apparatus and Materials

All stimuli were presented on a 30cm computer monitor in Times New Roman 24 font. The stimulus components were six symptoms; bad breath, stomach ache, skin rash blurred vision, ear ache, and nose bleed. A stimulus was presented in black letters, in the middle of one of four pink circles; one circle displayed in each of the corners of a black computer screen. On any given trial (compound or elemental), three stimuli were displayed, with one circle containing no stimulus (which circle was blank was randomly determined from trail to trail). This array was presented for 500ms. Following this array, one of two outcomes was presented in the center of a pink circle displayed in the middle of the computer screen; either the name of a fictitious disease (Jetson’s Syndrome or Hartley’s Disease) presented for 500ms, or no disease (in which case the computer screen remained blank for

500ms). There then followed a 500ms inter-trial interval, and another set of symptoms was presented.

Procedure

Each participant was tested individually in a small experimental room. Participants sat at a desk facing the experimenter and the computer, and were presented, on the screen, the instructions listed below:

"*You will be shown a series of slides. Each slide will show four discs, some discs will contain the name of a symptom. Some of these slides will be followed by an illness. Your task is to learn which symptoms predict the illnesses.*"

The participants were then presented with the task. Each block of training consisted of 48 presentations of the symptoms (48 symptom and outcome slides). The symptoms were designated to three sets: A+, BC+, DEF-; where A to F represent the symptoms, presented in random combinations across the participants, but always the same within-participant, and the + represents the two illnesses, again, randomly associated with either the element or compound across participants, but always the same within a participant. 'A' represents one symptom that always predicted one illness; BC represented two symptoms that always occurred together and always predicted the other illness, and D, E, and F, were three symptoms that did not reliably predict either illness. The other stimuli present on either A or BC trials were randomly selected from the DEF stimuli. On no illness trials, DEF were all present together. The precise symptoms represented by the letters differed between the participants randomly, but were always the same within-participant. The illnesses that were predicted by the elemental cue (A) or the compound (BC) were randomly assigned across participants, but were always associated with element or compound within a participant.

Participants were exposed to eight blocks of these trials (384 trials in total). At the end of each block of 48 trials, participants were presented with a list of the symptoms and the name of one of the diseases. They were asked to rate, on a scale of 0 (not predictive) to 100 (entirely predictive), each symptom with respect to that disease. Participants were then presented with the same list of symptoms and the name of the other disease. The order in which the foods were presented (top to bottom on the list) was random from presentation to presentation. The order in which the two allergies were asked about differed from block to block, and this order also differed between the participants.

Results and Discussion

Figure 1 about here

Figure 1 displays the mean ratings, across the eight blocks of training, for the elemental stimulus (A), the two stimuli from the compound (B and C – where B was the initially higher-rated stimulus of the pair on block 1), and for the mean rating for non-predictors (DEF). Inspection of these data shows that the element (A) was rated similarly to the initially-higher rated component of the compound (B), and A was also rated as more predictive than the other component of the compound (C). Although the two stimuli from the compound (B and C) were rated differently from one another at the start of training, this effect diminished as training progressed.

A two-factor repeated-measures analysis of variance (ANOVA) with (stimulus (A, B, C, and DEF), and block, as factors was conducted on these data. For this and all subsequent analyses, the appropriate Bayes statistic, for the alternative or null hypothesis, calculated using widely-adopted uninformative priors (Cauchy scale = 0.707), is also reported. This

ANOVA revealed statistically significant main effects of stimulus, $F(3,42) = 74.67$, $p < .001$, $\eta^2_p = .842$[95% *CI* = .729:.883], *Bayes Factor*, $p(H_1/D) = .999$, and block, $F(7,98) = 16.98$, $p < .001$, $\eta^2_p = .548$[.380:.619], $p(H_1/D) = .999$, and a significant interaction between the two factors, $F(21,294) = 5.92$, $p < .001$, $\eta^2_p = .297$[.142:.305], $p(H_1/D) = .744$. Simple effect analyses of stimulus on each block revealed a significant difference between the stimuli at each block, smallest $F(3,42) = 13.23$, $p < .001$. To further analyze the differences between the ratings given to the stimuli, paired t-tests (with Holm-Bonferroni corrections applied for each block to adjust the level of statistical significance), were conducted between each stimulus on each block. Stimuli A and B never differed from one another; Stimuli A and C differed on blocks 1 to 5, inclusively; and Stimuli A and D differed from one another on every block of training. Stimuli B and C differed from one another on blocks 1 to 4, inclusive; and Stimuli B and D differed on all blocks. Stimuli C and D differed from one another on blocks 3 to 8, inclusive.

To further isolate the impact of extended training on over-selectivity effects (i.e., the comparison of stimuli B and C), a two-factor ANOVA (stimulus x block) was conducted on the data from just these two stimuli. This revealed statistically significant main effects of stimulus, $F(1,14) = 47.30$, $p < .001$, $\eta^2_p = .772$[.441:.863], $p(H_1/D) = .999$, and block, $F(7,98) = 23.29$, $p < .001$, $\eta^2_p = .625$[.476:.685], $p(H_1/D) = .999$, and a significant interaction between the two factors, $F(7,98) = 3.60$, $p < .05$, $\eta^2_p = .205$[.036:.290], $p(H_1/D) = .999$. Simple effect analyses of stimulus on each block revealed a moderate significant difference between the stimuli at block 1, $F(1,98) = 23.60$, $p < .001$, $\eta^2_p = .194$[.071:.324], $p(H_1/D) = .627$, a large difference at block 2, $F(1,98) = 46.56$, $p < .001$, $\eta^2_p = .322$[.177:.447], $p(H_1/D) = .841$, a moderate difference at block 3, $F(1,98) = 22.99$, $p < .001$, $\eta^2_p = .190$[.069:.320], $p(H_1/D) = .572$, block 4, $F(1,98) = 26.62$, $p < .001$, $\eta^2_p = .214$[.086:.344], $p(H_1/D) = .641$, a small difference at block 5, $F(1,98) = 5.74$, $p < .05$, $\eta^2_p = .055$[.001:.161], $p(H_1/D) = .287$, block 6,

$F(1,98) = 9.83$, $p < .01$, $\eta^2_p = .119[.012:.209]$, $p(H_1/D) = .352$, but no difference at block 7, $F(1,98) = 2.76$, $p > .10$, $\eta^2_p = .028[.000:.117]$, $p(H_o/D) = .757$, or block 8, $F(1,98) = 4.11$, $p > .05$, $\eta^2_p = .040[000:.138]$, $p(H_o/D) = .737$.

These results show that ratings of the relationship between the symptoms and the illnesses developed over successive blocks of training. Moreover, there was an over-selectivity effect noted in this judgment procedure as has been noted previously in simultaneous discrimination procedures (e.g., Lovaas et al., 1971; Reed & Reynolds, 2011a) in that one element of the compound was rated consistently more highly than the other during the initial training period. The observed over-selectivity effect diminished with extended training. This finding is consistent with the effects previously noted for human over-selectivity (Gibson & Reed, 2005; Reynolds & Reed, 2011b), and also for nonhuman overshadowing effects (e.g., Bellingham & Gillette, 1981; Stout et al., 2003). It might be considered that the diminution of the over-selectivity effect noted here is due to a ceiling effect; that is, Stimulus B remains at ceiling while Stimulus C continues to be learned about. However, it might be noted that the reduction in the difference between B and C was noticeable by trial 5, and learning to all stimuli continued to increase beyond this point, making this suggestion less likely.

Experiment 2

Experiment 2 sought to further replicate these findings, but employed a between-subject version of Experiment 1. This was done to determine if the repeated judgments made by the participants in Experiment 1influenced the results (e.g., Catena, Maldonado, & Cándido, 1998). It has been suggested that, under some conditions, repeatedly judging the same events, rather than making one single judgment at the end of training, will affect the manner in which information is processed (see Catena et al., 1998; Miller & Matzel, 1987). To determine whether or not this would be the case, different groups of participants were required to make only one judgment after different amounts of training. If the effects noted in Experiment 1 were replicated in this study, then the above argument about repeated judgments impacting the results could not be sustained.

Experiment 2 also sought to assess whether there were any idiosyncratic biases in the participants' ratings of the likelihood of particular symptoms being associated with a particular illness. For example, it may be that a participant believes that nose bleeds are likely to be indicative of illness more than bad breath, and, if these symptoms are paired together, then this may explain the over-selectivity effect. Of course, given the randomization procedures employed, this would have to be true for many participants across a range of stimuli – but it is possible. To test this, the participants were asked to rate the likelihood of the symptoms predicting the two illnesses at the start of training before any exposure to the training procedure. Any differences in these ratings could then be related to those noted in the judgment to ascertain if those stimuli rated as more likely to be associated with an illness pre-training became the over-selected stimulus during training.

Method

Participants

Forty-eight participants (16 male and 32 female) aged between 18 and 30 years took part. None of the participants had any prior experience with the task, and all were volunteers recruited in the Department of Psychology. The apparatus and materials were as described in Experiment 1.

Procedure

The procedure was as described in Experiment 1 except that before training participants were presented with a list of the symptoms and the name of a disease, and were asked to rate on a scale of 0 (not predictive) to 100 (entirely predictive) each symptom, they were then presented with the same list and the name of the other disease (these were presented in random order across participants as described in Experiment 1). The participants were divided into four groups ($n = 12$). Each group received the same exposure to the stimuli (A+, BC+, DEF-), as noted in Experiment 1. Group 1 received only one block of 48 presentations of the stimuli, Group 2 received two blocks of 48 stimulus presentations, Group 4 received 4 blocks of 48 presentations, and Group 8 received 8 blocks of 48 trials. At the end of training, participants were presented with a list of the symptoms and the name of a disease, and were asked to rate on a scale of 0 (not predictive) to 100 (entirely predictive) each symptom as described above.

Results and Discussion

The pre-training ratings for the stimuli in terms of predicting the illness that was to be associated with the elemental cue were: A = 14.74 ($\pm$ 12.84); B = 14.16 ($\pm$ 12.84); C = 14.08 ($\pm$ 12.72); and DEF = 15.66 ($\pm$ 15.38); $F < 1$, $\eta^2_p = .002$. In terms of predicting the illness that was to be associated with the compound cue the pre-training means were: A = 15.00 ($\pm$ 15.28); B = 14.50 ($\pm$ 15.88); C = 18.70 ($\pm$ 19.26); and DEF = 20.88 ($\pm$ 15.24); $F(3,132) = 1.50$, $p > .20$, $\eta^2_p = .033$ (15 participants rated the symptom that was to become B higher than that which was to become C; and 21 rated C higher than B).

Figure 2 about here

Figure 2 displays the mean ratings for element (A), for the two stimuli from the compound (B and C), and the mean rating for the non-predictors (DEF), at the end of training, for all four groups. Inspection of these shows that as the amount of training prior to making a judgment increased, the ratings given to the stimuli increased. In general, increasing the amount of training produced a decrease in the level of over-selectivity (comparing B with C). Additionally, there was no difference between the rating given to the element (A) and the initially higher-rated component of the compound (B), but that A did differ from the other component C in the groups with lower levels of training.

A two-factor mixed-model ANOVA (group x stimulus) was conducted on these data and revealed statistically significant main effects of group, $F(3,44) = 51.50$, $p < .001$, $\eta^2_p = .778[.631:.845]$, $p(H_1/D) = .999$, and stimulus, $F(3,132) = 196.56$, $p < .001$, $\eta^2_p = .931[.904:.941]$, $p(H_1/D) = .999$, and a statistically significant interaction between these two factors, $F(9,132) = 14.24$, $p < .001$, $\eta^2_p = .493[.337:.558]$, $p(H_1/D) = .999$. Simple effect analyses revealed that there was a statistically significant main effect of stimulus for each group, smallest $F(3,132) = 12.28$, $p < .001$. Paired t-tests were conducted separately for each group: Stimulus A never differed from Stimulus B; it differed from Stimulus C in all but Group 8; and always differed from Stimulus DEF. Stimulus B differed from Stimulus C in all groups except Group 8; and it always differed from Stimulus DEF. Stimulus C differed from Stimulus DEF in Groups 6 and 8.

To further isolate the impact of extended training on over-selectivity effects (i.e. the comparison of stimuli B and C), a two-factor ANOVA (stimulus x block) was conducted on the data from just these two stimuli. This revealed statistically significant main effects of group, $F(3,44) = 41.34$, $p < .001$, $\eta^2_p = .738[.569:.805]$, $p(H_1/D) = .999$, and stimulus, $F(1,44) = 74.16$, $p < .001$, $\eta^2_p = .628[.433:.733]$, $p(H_1/D) = .999$, and a statistically significant interaction between these two factors, $F(3,44) = 9.64$, $p < .01$, $\eta^2_p = .400[.143:.537]$, $p(H_1/D)$

= .785. Simple effect analyses revealed that there was a large statistically significant main effect of stimulus for group 1, $F(1,44) = 16.95$, $p < .001$, $\eta^2_p = .278[.077:.458]$, $p(H_1/D) = .999$, and group 2, $F(1,44) = 17.85$, $p < .001$, $\eta^2_p = .289[.084:.467]$, $p(H_1/D) = .999$, a smaller effect for group 4, $F(1,44) = 5.41$, $p < .05$, $\eta^2_p = .109[.001:.290]$, $p(H_1/D) = .945$, but no effect for group 8, $F < 1$, $p > .90$, $\eta^2_p = .001[.000:.014]$, $p(H_o/D) = .829$.

These data show that the over-selectivity effect decreased as the groups received more training, even though the ratings were made at the end of training and not repeatedly. These findings replicate those reported in the current Experiment 1, and which have also been noted in previous studies of over-selectivity using different procedures (Lovaas et al., 1971; Reynolds & Reed, 2011b). The current study also demonstrated that this diminution of cue competition was not a product of the repeated judgments made about the stimuli in Experiment 1, but could also be noted when using a between-group design when only one judgment had been made. Of course, this is not to say that path-dependence effects do not occur in such procedures (e.g., see Catena et al., 1998), but they do not appear to impact the current effect. Given the range of studies that have reported such a finding, this may not be entirely surprising. The diminution effect was also apparent before conditioning was at asymptote, with the difference between the ratings for stimuli B and C being smaller in Group 4 than in Groups 1 and 2, but while there was still room for improvement in ratings (compare with Group 8). This suggests that ceiling effects were not entirely responsible for these effects. Furthermore, there was evidence that any pre-existing views about the likelihood of a particular symptom being associated with a particular illness were not responsible for the over-selectivity effects, which appear to occur during training itself.

Experiment 3

The preceding experiments have demonstrated that an over-selectivity effect can be obtained in a human causal judgment procedure, and that this effect reduces in size as training continues. This finding is in line with previous findings for overshadowing effects (Bellingham & Gillette, 1981; Stout et al., 2003). This suggests that there are potential similarities between over-selectivity and other cue competition effects. Given this, it may be that explanations developed for explaining overshadowing phenomena also could be applied to over-selectivity. Of particular significance in this latter regard has been the role of within-compound associations. It has been suggested that when strong within-compound associations are formed, such as with extended training (Stout et al., 2003), long duration CS presentations (Sissons et al., 2009), or stimuli that are spatially contiguous (Glautier, 2002), then overshadowing will be less pronounced. This may also apply to the size of the over-selectivity effect in the current experiments.

Given this, the current experiment sought to establish whether greater levels of training would lead to less over-selectivity, and whether this is might be associated with stronger within-compound associations. To this end, participants were exposed to rating task as described above for shorter or longer durations, but with a modification from the preceding studies. The participants were given a test to measure their ability to remember compounds actually presented during the initial training phase (Wasserman & Berglan, 1998), and their confidence in that judgment (Luque, Flores, & Vadillo, 2015). To the extent that they performed well on this task, they might be regarded as having established strong within-compound associations. This score should correlate with the relative absence of over-selectivity, and the degree of over-selectivity should be seen to vary as this measure of the within-compound strength varies (even irrespective of the actual training procedure adopted, see Luque et al., 2015; Wasserman & Berglan, 1998).

Method

Participants

Sixty-four participants (20 male and 44 female) aged between 19 and 42 years took part. None of the participants had any prior experience with the task, and all were volunteers recruited in the Department of Psychology. The apparatus and materials were as described in Experiment 2.

Procedure

The procedure was as described in Experiment 2 except participants were divided into two groups (*n* = 32): Group 2 received two blocks of 48 stimulus presentations (containing A+ BC+, and DEF-) stimuli; and Group 8 received 8 blocks of 48 presentations. After giving their final ratings concerning the symptoms, the recognition test was presented to all participants. Participants had to select the compounds that had been presented during the initial training. To this end, participants were presented with a pencil and paper test, in which 15 pairs of stimuli were created from the 6 elements (a combination of each stimulus with every other stimulus). The participants were asked to circle any combinations that they had seen before, and to rate their confidence in their choices on a scale of 0 to 9, in which a score of 0 indicated being “completely unconfident”, and a score of 9 indicated being “absolutely confident”. The order of the combinations was random across participants.

Results and Discussion

The pre-training ratings for the stimuli in terms of predicting the illness that was to be associated with the elemental cue were: A = 9.92 ($\pm$ 7.94); B = 9.59 ($\pm$ 7.43); C = 10.39 ($\pm$ 7.93); and DEF = 12.66 ($\pm$ 7.61); $F(3,189) = 1.57$, p > .20, $\eta^2_p = .197$. In terms of predicting the illness that was to be associated with the compound cue the pre-training means were: A = 9.64 ($\pm$ 7.13); B = 10.20 ($\pm$ 7.95); C = 9.48 ($\pm$ 8.76); and DEF = 12.24 ($\pm$ 6.07); $F(3,189) = 1.86$, $p > .10$, $\eta^2_p = .137$ (28 participants rated the symptom that was to become B higher than that which was to become C; and 24 rated C higher than B).

Figure 3 about here

Figure 3 displays the mean ratings for element (A), for the two stimuli from the compound (B and C), and the mean rating for the non-predictors (DEF), at the end of training for the groups with 2 and 8 blocks of training. Inspection of these shows that as the amount of training prior to making a judgment increased, the ratings given to the stimuli increased. In general, increasing the amount of training produced a decrease in the level of over-selectivity (comparing B with C). Additionally, there was no difference between the rating given to the element (A) and the initially higher-rated component of the compound (B), but that A did differ from the other component C in the groups with lower levels of training.

A two-factor mixed-model ANOVA (group x stimulus) was conducted on these data and revealed statistically significant main effects of group, $F(1,62) = 54.95$, $p < .001$, $\eta^2_p = .470[.284:.597]$, $p(H_1/D) = .999$, and stimulus, $F(3,186) = 814.32$, $p < .001$, $\eta^2_p = .929[.911:.941]$, $p(H_1/D) = .999$, and a statistically significant interaction between these two factors, $F(3,186) = 53.84$, $p < .001$, $\eta^2_p = .465[.357:.541]$, $p(H_1/D) = .999$. Simple effect analyses revealed that there was a statistically significant main effect of stimulus for each group, smallest $F(3,168) = 237.49$, $p < .001$.

To further isolate the impact of extended training on over-selectivity effects (i.e. the comparison of stimuli B and C), a two-factor ANOVA (stimulus x block) was conducted on the data from just these two stimuli. This revealed a statistically significant main effects of group, $F(1,62) = 78.89$, $p < .001$, $\eta^2_p = .560[.387:.669]$, $p(H_1/D) = .999$, and stimulus, $F(1,62) = 129.84$, $p < .001$, $\eta^2_p = .677[.534:.758]$, $p(H_1/D) = .999$, and a statistically significant interaction between these two factors, $F(1,62) = 81.13$, $p < .001$, $\eta^2_p = .567[.395:.674]$, $p(H_1/D) = .999$. Simple effect analyses revealed that there was a large-sized statistically significant main effect of stimulus for group 2, $F(1,62) = 216.$, $p < .001$, $\eta^2_p = .777[.671:.834]$, $p(H_1/D) = .999$, but no significant effect for group 8, $F(1,62) = 2.42$, $p > .10$, $\eta^2_p = .038[.000:.163]$, $p(H_o/D) = .734$.

Figure 4 about here

The confidence ratings given to the pairs of stimuli were noted, and the mean ratings given to the target and the pseudo compounds were calculated for each group. The mean confidence rating for the target compound for Group 2 was 5.84 ($\pm$ 1.61), and for Group 8 this was 7.81 ($\pm$ .97), $t(32) = 5.94$, $p < .001$, $d = 1.52$. The mean confidence rating for the non-target compounds for Group 2 was 3.45 ($\pm$ .98), and for Group 8 this was 2.69 ($\pm$ .90), $t(32) = 3.19$, $p < .01$, $d = .79$. The mean difference in the confidence ratings between the target and non-target compounds for Group 2 was 2.41 ($\pm$ 1.90), and for Group 8 this was 5.13 ($\pm$ 1.43), $t(32) = 6.47$, $p < .001$, $d = 1.64$. Correlational analyses conducted on these data revealed statistically significant negative correlations between the level of over-selectivity (rating for B minus rating for C) and the confidence ratings for the target compound, $r(62) = -.694$, $p < .001$. There was a statistically significant positive correlation between the level of over-selectivity and the confidence ratings for the non-target compounds, $r(62) = .254$, $p < .05$, but a negative correlation between the difference score and over-selectivity, $r(62) = -.648$, $p < .001$. Figure 4 shows the scatterplots for the relationships between confidence (top panel = target; middle panel = nontargetr; bottom panel = target minus nontarget) and levels of over-selectivity (most minus least-selected stimulus).

Figure 5 about here

The sample was also subject to a mean split (sample mean = 6.82 $\pm$ 1.64) in terms of their confidence ratings in recognizing the actual compound presented during training. Those with a score of 7 or above were classified as having higher confidence. This produced a low-confidence group (n = 25, Group 2 = 21, Group 8 = 4; mean = 5.08 + 1.00; range = 3 – 6), and a high-confidence group (n = 39, Group 2 = 11, Group 28 = 4; mean = 7.95 $\pm$ .76; range = 7 – 9). The ratings for stimuli B and C for these two groups are show in Figure 5, and reveal a much greater difference between these stimuli for the low-confidence group compared to the high-confidence group. This observation was corroborated by a two-factor mixed-model ANOVA (group x stimulus), which revealed statistically significant main effects of group, $F(1,62) = 22.08$, $p < .001$, $\eta^2_p = .263[.093:.420]$, $p(H_1/D) = .999$, and stimulus, $F(1,62) = 114.08$, $p < .001$, $\eta^2_p = .648[.497:.736]$, $p(H_1/D) = .999$, and a statistically significant interaction between these two factors, $F(1,62) = 36.01$, $p < .001$, $\eta^2_p = .367[.181:.513]$, $p(H_1/D) = .999$. Simple effect analyses revealed that there was a large-sized statistically significant main effect of stimulus for low confidence group, $F(1,62) = 21.62$, $p < .001$, $\eta^2_p = .259[.090:.416]$, $p(H_1/D) = .999$, but no significant effect for the high confidence group, $F(1,62) = 2.37$, $p > .10$, $\eta^2_p = .037[.000:.162]$, $p(H_o/D) = .708$.

Taken together, these data replicate the previous findings reported here, that over-selectivity reduces with extended training, and they suggest that this effect is similar to that seen for other cue competition paradigms (e.g., Stout et al., 2003). They also support the view that extended training works to increase the strength of the within-compound association (see also Luque et al., 2015); such that, the stronger is the within-compound association, the weaker is the degree of over-selectivity. Moreover, the strength of the within-compound association, irrespective of actual assignment to training group, also predicts the level of over-selectivity seen. An effect also noted previously for unovershadowing and backward blocking effects (Luque et al., 2015). As with Experiment 2, there was nothing in the pre-training ratings that implies that these judgments were merely the result of pre-existing basis in the participants.

General Discussion

The current studies explored the degree to which stimulus over-selectivity could be observed in a human judgment procedure, and also to determine the impact of different levels of training on this effect. The results from all three experiments demonstrate that, when two stimuli were presented in compound prior to an outcome, one of these stimuli would be rated as more related to the outcome than the other, despite the stimuli having a similar predictive validity with respect to that outcome. These findings mirror those produced in studies of over-selectivity using a concurrent discrimination procedure (e.g., Lovaas et al., 1971; Reed et al., 2009; Reed & Gibson, 2005), in which the elements of a compound stimulus (AB+), which are of equal predictive validity to one another, are differentially effective in controlling behavior when presented separately from one another.

This over-selectivity effect is similar to a unilateral overshadowing effect, which has been noted in studies of animal conditioning (see Mackintosh, 1976; Stout et al., 2003). Moreover, the current over-selectivity effect decreased with increased levels of training in all studies, irrespective of whether there were multiple judgments (Experiment 1) or one judgment at the end of different amounts of training (Experiments 2 and 3). A similar impact of extended training has been observed in studies of over-selectivity using a simultaneous discrimination procedure with humans (Reynolds & Reed, 2011b), and in studies of overshadowing using nonhumans (e.g., Bellingham & Gillette, 1981; Stout et al., 2003); that is, as training proceeds, the level of overshadowing between two stimuli presented in compound diminishes.

In terms of overshadowing, a number of views have been put forward as to why increased levels of training might decrease the observed level of overshadowing. It has been suggested that as the strength of the within-compound association grows, as with extended training, overshadowing will decrease. The current Experiment 3 extended this view to the current over-selectivity findings, and noted that when measures of strong within-compound associations were high (see Wasserman & Berglan, 1998; Luque et al., 2015), that over-selectivity was low. Extended training is not the only manipulation that might produce such an effect; long duration CS presentations (Sissons et al., 2009), and stimuli that are spatially contiguous (Glautier, 2002), and these could also be examined in the context of over-selectivity. Such a view may also explain the impact of a range of other factors on over-selectivity – all of which seem to show that when learning is strong that over-selectivity is weak. For example, the effect occurs more readily when employing partial as opposed to continuous reinforcement schedules (Reynolds & Reed, 2011b), following trace as opposed to delayed conditioning procedures (Gibson & Reed, 2005), and with groups known to have difficulty with forming within-compound associations, such as those with ASD (see Plaisted, O'Riordan, & Baron-Cohen, 1998; Reed, 2011). This suggests that consideration of the role of within-compound associations will also be important in explaining over-selectivity effects, and brings this phenomenon into an associative framework.

A number of models of associative learning could be employed to explain the current findings for over-selectivity (Denniston et al., 2001; Mackintosh, 1976; Pineno, 2007). Pineno (2007) suggested a cue facilitation model that assumes that AB+ trials result in competition between A and B for associative strength in the way described by Rescorla and Wagner (1972). In addition, A-B and B-A within-compound associations are learned. At test of A, excitatory responding is increased by the representation of the outcome activated through the A-outcome, and the A-B-outcome associative chain. However, generalisation of this model to the current context might be limited by procedural differences during training and test, and by the fact that it assumes an important role for novelty. Specifically, that the A-B-outcome associative chain increases responding to A, only to the extent that A is novel, which would not account for the effects of extended training.

Predictions derived from some associability theories (e.g., Mackintosh, 1975) suggest that overshadowing (and potentially over-selectivity) would be greater at lower levels of conditioning, which would not be the case after extended training (see Mackintosh, 1976, Stout et al., 2003, for discussions). However, these views have difficulty in accommodating the revaluation studies in over-selectivity contexts (see Reed et al., 2009; Reed et al., 2012). In these studies, a AB+ CD- simultaneous discrimination procedure was adopted, and then participants were tested non-reinforced for their choices between the elements (i.e. AvC, AvD, BvC, BvD), and one of the previously reinforced elements (e.g., A) was chosen more often than the other (B). In a subsequent phase, the most selected stimulus from the previously reinforced pair (i.e. A or B) was presented non-reinforced in a simple discrimination training procedure (i.e. A- E+), and, when re-tested (AvC, AvD, BvC, BvD), not only was element A picked less, but the previously under-selected stimulus (B) was chosen more often despite having no direct conditioning.

Denniston et al. (2001) developed the comparator hypothesis (Miller & Matzel, 1998) that could explain these results. Within comparator models, contiguity is assumed to be the only requisite for learning to occur (see Denniston et al, 2001; Miller & Matzel, 1988); thus, it is anticipated that on any target-outcome trial, all of the stimuli present will acquire some degree of strength. In order to account for overshadowing, focus is centered on the manner in which learning is manifested, or expressed, during the testing process. In particular, the expression of learning is assumed to be a direct result of a CS(A)-US association, and an inverse function of a CS(A)-CS(X) association and a CS(X)-US association; where CS(A) is the 'target' stimulus, and CS(X) is the 'comparator' stimulus (see Miller & Matzel, 1988). The model assumes that the co-occurrence of the comparator stimulus, CS(X), will have no impact on the learning accrued to the target, CS(A), but the expression of learning is impaired during the testing process when a comparison is made between the association strengths of CS(A) and CS(X).

Denniston et al. (2001) proposed that a sufficient amount of training could result in the context (i.e., the environmental features of the context in which conditioning occurs) acting as 'second-order comparator stimuli' to the initial 'comparator stimuli', now denoted the 'first order comparator stimuli'. According to this 'extended' comparator hypothesis, it is postulated that the emergence of a second-order comparator could reduce the initial 'first order' comparator's ability to attenuate the 'target' CS(A)-US association. Within overshadowing and over-selectivity procedures, it could be that the conditioning context, acting as the second-order comparator, attenuates the ability of the comparator stimulus, CS(X), to weaken the response to CS(A). This suggests that the overshadowing effect would dissipate as a result of prolonged training, a result which would directly contrast the predictions of the Rescorla-Wagner model. It should be noted that in order to accommodate the current length of training effects they would need to suggest that significant conditioning occurred to the context, and it is unclear precisely what 'context' would refer to in such human judgment procedures. Moreover, according to the version of the comparator theory articulated by Denniston et al. (2001), while context conditioning should reduce the response to B and increase the response to C – but the former effect did not happen in the present experiments. There are also a range of theories derived to accommodate the findings from human judgment studies, but which do not rely on associative assumptions (e.g., Cheng, 1997; De Houwer, Beckers, & Glautier, 2002; White, 2005). Although these views may be made to explain the current results, it is difficult to see their application to the results of over-selectivity studies conducted using the simultaneous compound discrimination procedures.

Of course, there are a number of limitations to the current study that should be acknowledged. The current learning task is different from previously employed associative learning tasks (e.g., Van Hamme & Wasserman, 1994). For example, it is not common to present the stimuli for just 500ms at the corners of the screen. This procedure has been used previously for examination of judgements (Quigley & Reed, 2017), and does allow the advantage of extending generality of the findings, but replication with a more standard procedure would seem warranted.

It is also worth noting that in a range of over-selectivity studies, as described in the General Introduction, the difference in the extent to which the elements of a compound stimulus control behavior decreases as learning about the target becomes stronger (see Gibson & Reed, 2005; Reed, 2011; Reed & Gibson, 2005; Reynolds & Reed, 2011b). The stimulus over-selectivity effect has been shown to occur more strongly in a range of clinical populations; most often with children with autism spectrum disorder (Leader, Loughnane, Mc Moreland, & Reed, 2009; Lovaas et al., 1971; Reed, Broomfield, McHugh, McCausland, &

Leader, 2009) and learning difficulties (Dube et al., 2009; Dube & McIlvane, 1999), but also with people with acquired brain injury (Wayland & Taplin, 1985), and older people (McHugh & Reed, 2007). Recently, over-selectivity has been shown in populations lacking any clinical disorder, but who are under a degree of cognitive strain, produced by performing a concurrent task (Reed & Gibson, 2005; Reed et al., 2012; Reynolds & Reed, 2011a), and this effect of limited stimulus control accruing to one element of a stimulus in situations requiring high cognitive demands is similar to that noted in a number of attention-based tasks (e.g., Kim, Kim, & Chun, 2005).

Thus, the currents result show that unilateral overshadowing/over-selectivity effects can be seen in human judgment studies, and that these effects dissipate with extended training, as do overshadowing effects. This strengthens the link between these cue competition effects, but also produces some difficulty in finding a common explanation that can be applied to accommodate these effects.

References

Azorlosa, J.L., & Cicala, G.A. (1988). Increased conditioning in rats to a blocked CS after the first compound trial. *Bulletin of the Psychonomic Society*, **26**, 254-257.

Bellingham, W.P., & Gillette, K. (1981). Attenuation of overshadowing as a function of nondifferential compound conditioning trials. *Bulletin of the Psychonomic Society*, **18**, 218-220.

Brown S., & Bebko J.M. (2012). Generalization, overselectivity, and discrimination in the autism phenotype: a review. *Research in Autism Spectrum Disorders*. **6**, 733–740.

Catena, A., Maldonado, A., & Cándido, A. (1998). The effect of frequency of judgement and the type of trials on covariation learning. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(2), 481.

Chapman, G.B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Mem*ory, &: *Cognition*, **17**, 837-854.

Cheng, P.W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, **104**, 367–405.

De Houwer, J., Beckers, T., & Glautier, S. (2002). Outcome and cue properties modulate blocking. *Quarterly Journal of Experimental Psychology*, **55A**, 965–985.

Denniston, J. C., Savastano, H. I., & Miller, R. R. (2001). The extended comparator hypothesis: Learning by contiguity, responding by relative strength. In R. Mowrer & S. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 65-117). Hillsdale, NJ: Erlbaum.

Dickinson, A., & Burke, J. (1996). Within compound associations mediate the retrospective revaluation of causality judgements. *Quarterly Journal of Experimental Psychology*, **49**, 60-80.

Dube, W. V. (2009). Stimulus overselectivity in discrimination learning. In P. Reed (Ed.), *Behavioral theories and interventions for autism*, 23-46. New York: Nova.

Dube, W.V., Lombard, K.M., Farren, K.M., Flusser, D., Balsamo, L.M., & Fowler, T.R. (1999). Eye tracking assessment of stimulus overselectivity in individuals with mental retardation. *Experimental Analysis of Human Behaviour Bulletin*, **13**, 267-271.

Dube, W.V. & McIlvane, W. J. (1999). Reduction of stimulus overselectivity with nonverbal differential observing responses. *Journal of Applied Behaviour Analysis*, **32**, 25-33.

Gibson, E., & Reed, P. (2005). Stimulus over-selectivity in rats. *Journal of Autism and Developmental Disorders*, **35**, 851-859.

Glautier, S. (2002). Spatial separation of target and competitor cues enhances blocking of human causality judgements. *The Quarterly Journal of Experimental Psychology: Section B*, **55**, 121-135.

James, J.H., & Wagner, A.R. (1980). One-trial overshadowing: evidence of distributive processing. *Journal of Experimental Psychology: Animal Behavior Processes*, **6**, 188-205.

Kamin L.J. (1968). Attention-like" processes in classical conditioning. In M.R. Jones (ed.), *Miami Symposium on the Prediction of Behavior: Aversive Stimulation* (pp. 9–31). Miami, FL: Univ. Miami Press.

Kamin, L.J. (1969). Predictability, surprise, attention, and conditioning. *Punishment and aversive behavior*, 279-296.

Kaufman, M. A., & Bolles, R. C. (1981). A nonassociative aspect of overshadowing.

Bulletin of the Psychonomic Society, **18**, 318-320.

Kim, S.Y., Kim, M.S., & Chun, M.M. (2005). Concurrent memory load can reduce distraction. *Proceedings of the National Academy of Sciences*, **102**, 16524-16529.

Koegel, R.L., & Scheribman, L. (1977). Teaching autistic children to respond to simultaneous multiple cues. *Journal of Experimental Child Psychology*, **24**, 299-311.

Leader, G., Loughnane, A., Mc Moreland, C., & Reed, P. (2009). The effect of stimulus salience on over-selectivity. *Journal of Autism and Developmental Disorders*, **39**, 330-338.

Lovaas, O. I., Berberich, J. P., Perloff, B. F., & Schaeffer, B. (1966). Acquisition of imitative speech in schizophrenic children. *Science*, **151**, 705–707.

Lovaas, O. I. & Schreibman, L. (1971) Stimulus over selectivity of autistic children in a two stimulus situation. *Behaviour, Research and Therapy*, **9**, 305-310.

Lovaas, O. I., Schreibman, L., Koegel, R., & Rehm, R. (1971). Selective responding by autistic children to multiple sensory input. *Journal of Abnormal Psychology*, **77**, 211–222.

Luque, D., Flores, A., & Vadillo, M. A. (2013). Revisiting the role of within-compound associations in cue-interaction phenomena. *Learning & Behavior*, *41*(1), 61-76.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, **82**, 276-298.

Mackintosh**,** N.J. (1976)**.** Overshadowing and stimulus intensity. *Animal Learning and Behaviour*, **4**, 186-192.

Maes, E., Boddez, Y., Alfei, J. M., Krypotos, A. M., D'Hooge, R., De Houwer, J., & Beckers, T. (2016). The elusive nature of the blocking effect: 15 failures to replicate. *Journal* *of Experimental Psychology: General*, *145*(9), e49.

McHugh, L., & Reed, P. (2007). Age trends in stimulus over-selectivity. *Journal of the Experimental Analysis of Behavior*, **88**, 369-380.

Miller, R.R., & Matzel, L.D. (1987). Memory for associative history of a conditioned stimulus. *Learning& Motivation*, **18**, 118-130.

Miller, R.R., & Matzel, L.D. (1988). The comparator hypothesis: A response rule for the expression of associations. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 1-92). Orlando, FL: Academic Press.

Mitchell, C.J., Killedar, A., & Lovibond, P.F. (2005). Inference-based retrospective revaluation in human causal judgments requires knowledge of within compound relationships. *Journal of Experimental Psychology: Animal Behavior Processes, 31*(4), 418-424.

Pavlov, I. P. (1927). *Conditioned reflexes* (G. V. Anrep, Trans.). London: Oxford University Press.

Pearce, J.M., & Bouton, M.E. (2001). Theories of associative learning in animals. *Annual Review of Psychology*, **52**, 111–39.

Pineño, O. (2007). A response rule for positive and negative stimulus interaction in associative learning and performance. *Psychonomic bulletin & review, 14*(6), 1115-1124.

Plaisted, K., O'Riordan, M., & Baron-Cohen, S. (1998). Enhanced visual search for a conjunctive target in autism: A research note. *Journal of Child Psychology and Psychiatry*, **39**, 777-783.

Ploog, B.O. (2010). Stimulus overselectivity four decades later: A review of the literature and its implications for current research in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, **40**, 1332-1349.

Price, P.C., & Yates, J.F. (1993). Judgmental overshadowing: Further evidence of cue interaction in contingency judgment. *Memory* & *Cognition*, **21**, 561-572.

Quigley, M., & Reed, P. (2017). Overselective responding in a diagnostic judgment task. *Applied Cognitive Psychology,* **31**, 558-564.

Reed, P. (2011). Comparator mechanisms and Autistic Spectrum Conditions. In. T.R. Schachtman & S.R. Reilly (Eds.), *Associative Learning and Conditioning: Human and Animal Applications*. Oxford University Press.

Reed, P., Broomfield, L., McHugh, L., McCausland, A., & Leader, G. (2009). Extinction of over-selected stimuli causes re-emergence of previously under selected stimuli in higher functioning children with autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, **39**, 290-298.

Reed, P. & Gibson, E. (2005). The effects of concurrent task load on stimulus over selectivity. *Journal of Autism and Developmental Disorders,* **35**, 601-614.

Reed, P., Hawthorn, R., Bolger, S., Meredith, K. & Bishop, R. (2012). Disrupted stimulus control but not reward sensitivity in individuals with Autism Spectrum Disorders: A Matching Law Analysis. *Journal of Autism and Developmental Disorders* 42(11), 2393-2403.

Reed, P., Reynolds, G., & Fermandel, L. (2012). Revaluation procedures produce emergence of under-selected stimuli following simultaneous discrimination in humans. *Quarterly Journal of Experimental Psychology*.

Reed, P., Staytom, L., Stott, S. & Truzoli, R. (2011). Comparison of conditioning impairments in children with Down syndrome, autistic spectrum disorders and mental age-matched controls. *Journal of Intellectual Disability Research* 55(10), 988-997.

Rescorla, R.A., & Wagner, A. R. (1972).A theory of Pavlovian conditioning:

Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black &W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.

Reynolds, G., & Reed, P. (2011a). The strength and generality of stimulus over selectivity in simultaneous discrimination procedures. *Learning and Motivation*, **42**, 113-122.

Reynolds, G., & Reed, P. (2011b). The effects of schedule of reinforcement on over selectivity. *Research in Developmental Disabilities*, **32**, 2489-2501.

Schnur, P. (1971). Selective attention: Effect of element preexposure on compound conditioning in rats. *Journal of Comparative and Physiological Psychology*, **76**, 123-130.

Sissons, H.T., Urcelay, G.P., & Miller, R.R. (2009). Overshadowing and CS duration: Counteraction and a reexamination of the role of within-compound associations in cue competition. *Learning & Behavior*, **37**, 254-268.

Stout, S., Arcediano, F., Escobar, M., & Miller, R.R. (2003). Overshadowing as a function of trial number: Dynamics of first- and second-order comparator effects *Learning& Behavior*, **31**, *85-97*

Urcelay, G. P. (2017). Competition and facilitation in compound conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition*, *43*(4), 303.

Van Hamme, L.J., & Wasserman, E.A. (1994). Cue competitions in causality judgments: The role of nonpresentation of compound stimulus elements *Learning and Motivation*, **25**, 127-151.

Wagner, A.R., Logan, F.A., Haberlandt, K., & Price, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Comparative and Physiological Psychology*, **76**, 171–80.

Waldmann, M.R. (2001). Predictive versus diagnostic causal learning: Evidence from an overshadowing paradigm. *Psychonomic Bulletin & Review,* **8**, *600-608*.

Wasserman, E., & Berglan, L. R. (1998). Backward blocking and recovery from overshadowing in human causal judgement: The role of within-compound associations. *The Quarterly Journal of Experimental Psychology,* **51B**, 121-138.

Wasserman, E.A., & Miller, R.R. (1997). What's elementary about associative learning? *Annual Review of Psychology,* **48**, 573-607.

Wayland, S., & Taplin, J. E. (1985). Feature processing deficits following brain injury over selectivity in recognition memory for compound stimuli. *Brain and Cognition*, **4**, 338–355.

Wheeler, D.S., & Miller, R.R. (2008). Determinants of cue interactions. *Behavioural Processes*, *78*(2), 191-203.

White, P.A. (2005). Cue interaction effects in causal judgement: An interpretation in terms of the evidential evaluation model. *Quarterly Journal of Experimental Psychology*, **58B**, 99–140

Acknowledgements

Thanks are due to Lisa A. Osborne for her support. Correspondence regarding this article should be addressed to: Phil Reed, Department of Psychology, Swansea University, Singleton Park, Swansea, SA2 8PP, U.K. (e-mail: p.reed@swansea.ac.uk).

Figure 1. Results from Experiment 1. Mean judgments for stimuli after each block of training: A = elemental, B = initially higher-rated compound component, C = initially lower-rated compound component, DEF = nontarget stimuli. Error bars are not shown as comparisons are within-subject.

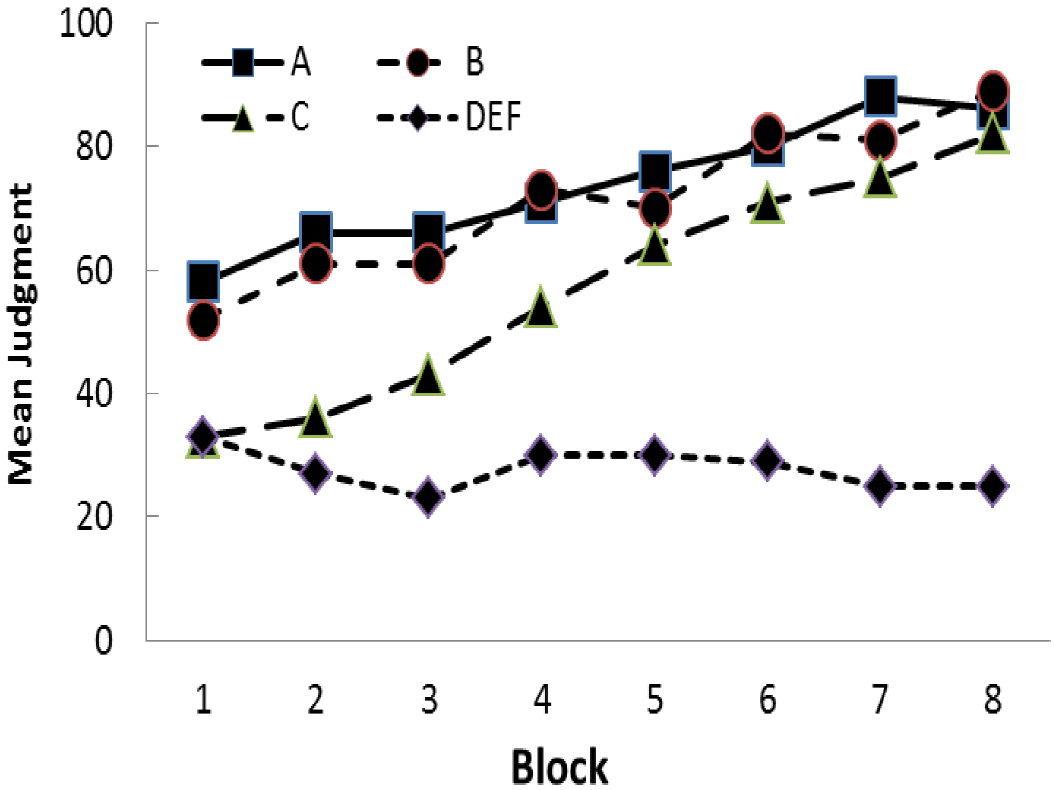

Figure 2. Results from Experiment 2. Mean judgments for stimuli for each group. Group 1 received 1 block of training; Group 2 received 2 blocks of training; Group 4 received 4 blocks of training; Group 8 received 8 blocks of training. A = elemental, B = initially higher-rated compound component, C = initially lower-rated compound component, DEF = nontarget stimuli. Error bars = 95% confidence intervals.

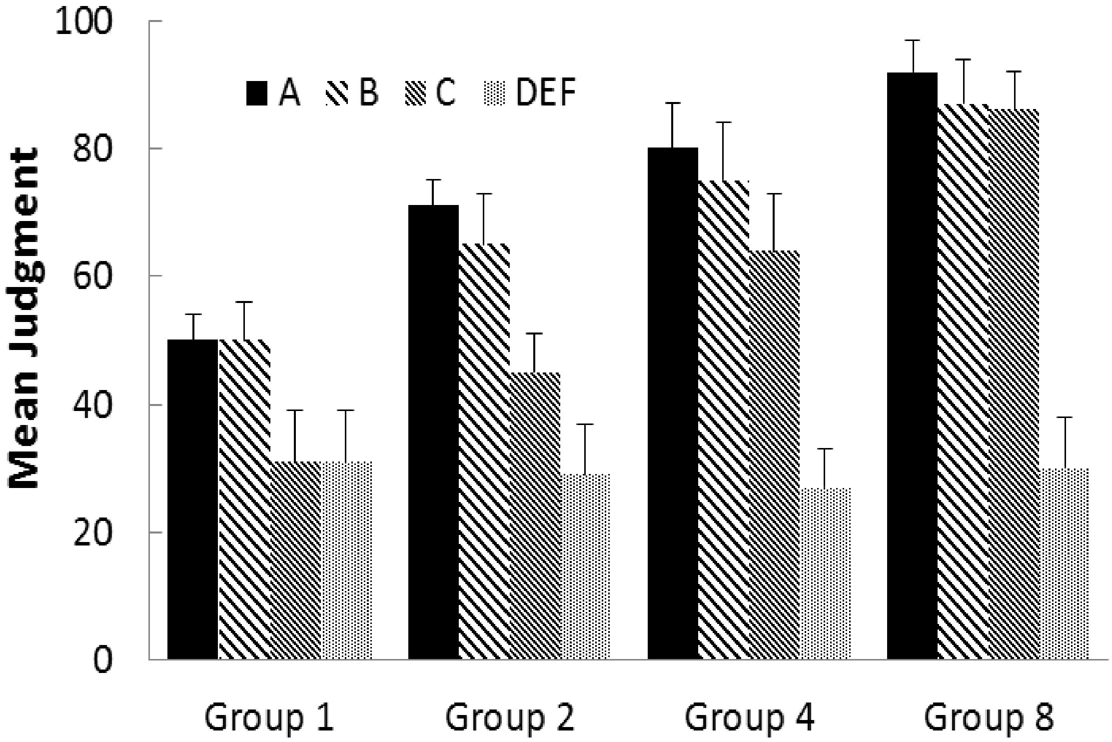

Figure 3. Results from Experiment 3. Mean judgments for stimuli for each group. Group 2 received 2 blocks of training; Group 8 received 8 blocks of training. A = elemental, B = initially higher-rated compound component, C = initially lower-rated compound component, DEF = nontarget stimuli. Error bars = 95% confidence intervals.

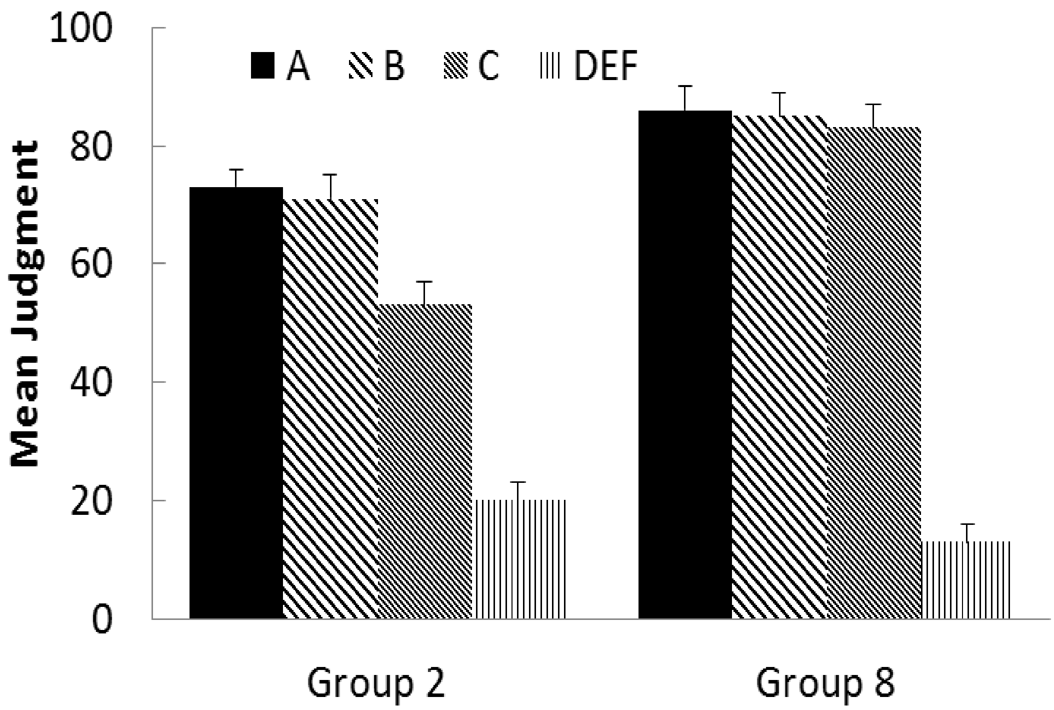

Figure 4. Experiment 3. Scatterplots showing relationships between confidence (top panel = target; middle panel = nontarget; bottom panel = target minus nontarget) and levels of over-selectivity (most minus least-selected stimulus).

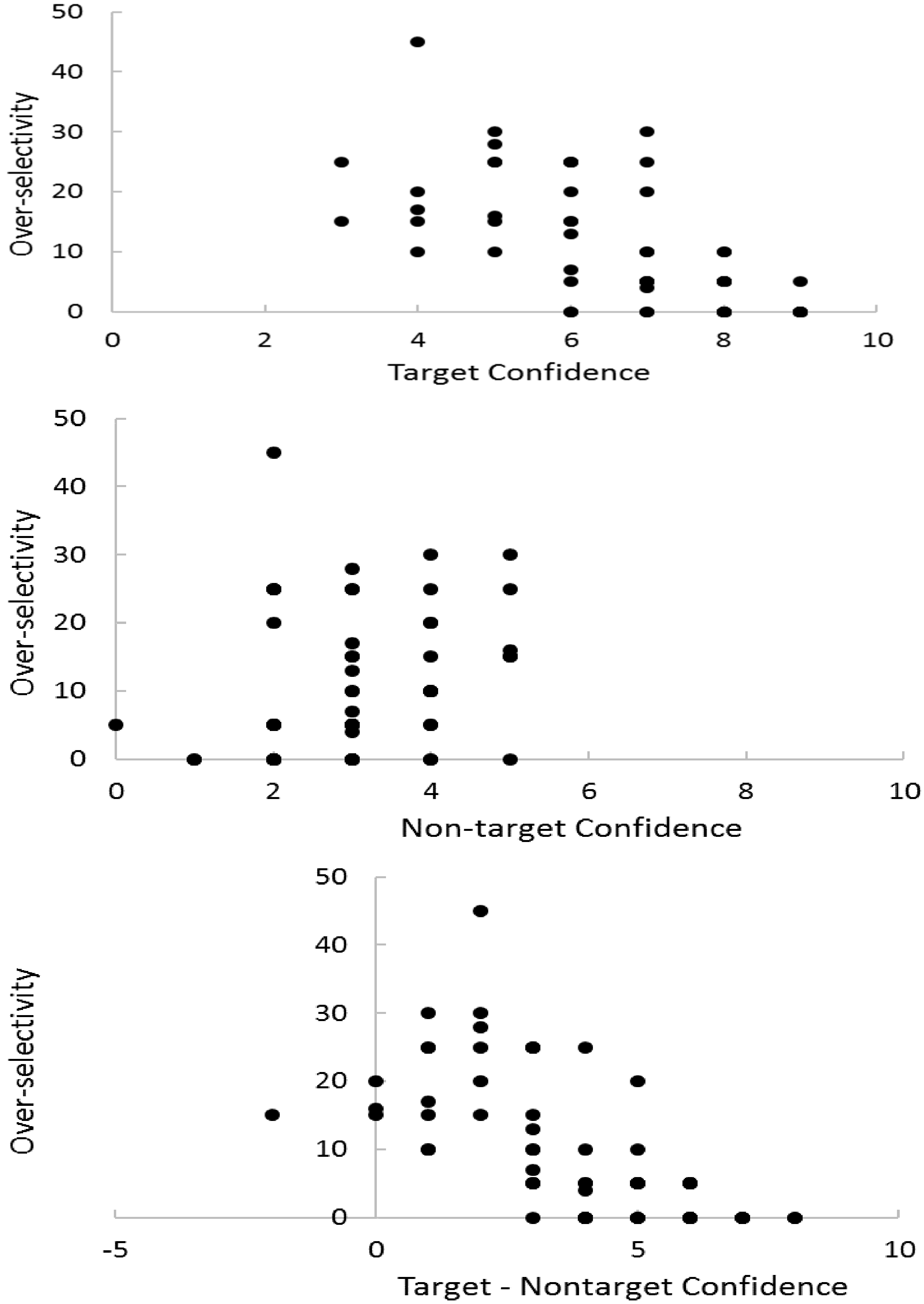

Figure 5. Results from Experiment 3. Mean judgments for stimuli for the high- and low-confidence group. B = initially higher-rated compound component, C = initially lower-rated compound component. Error bars = 95% confidence intervals.