Revaluing overselected stimuli: Effects of degree of posttraining extinction on stimulus overselectivity

Stephanie Gomes-Ng[1,2], Sarah Cowie[2], and Douglas Elliffe[2]

[1] Auckland University of Technology, Auckland, New Zealand

[2] The University of Auckland, Auckland, New Zealand

Correspondence

Stephanie Gomes-Ng, Department of Psychology, Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand.

Email: stef.gomes-ng@aut.ac.nz

Editor-in-Chief: Suzanne H. Mitchell

Academic Editor: Cynthia J. Pietras

Abstract

When responding to a stimulus exerting overselective control over behavior is extinguished, control by underselected stimuli may emerge. We investigated how the degree of extinction influences control by underselected stimuli. Adult humans ($N = 459$) chose between rapidly presented compound S+ and S- stimuli in a simultaneous discrimination. Then, participants chose between individual compound-stimulus elements in an unreinforced testing phase. The S+ element that was chosen most often underwent revaluation, during which choice of that element was reinforced with a probability ranging from 0 (complete extinction) to 1 (no extinction) in different groups. In post-revaluation retesting, choice of the overselected element was lower than pre-revaluation testing; this decrease was greater when the overselected element had been reinforced with a lower probability during revaluation. For the underselected element, choice decreased when the overselected element was completely extinguished and increased when the overselected element was sometimes or always reinforced. This highlights the role of the contingency change in post-revaluation changes in stimulus control. Our findings are consistent with comparator theories of overselectivity and suggest that control by underselected stimuli may emerge after partial extinction of an overselected stimulus. Future studies should establish the generality of these findings with clinical populations displaying overselectivity.

KEYWORDS

compound stimulus, extinction, humans, retrospective revaluation, stimulus overselectivity

The presence of multiple stimuli signaling the consequences that will follow a behavior does not guarantee that all of those stimuli will exert equal behavioral control (Skinner, 1953). In some situations, only one stimulus or a small subset of stimuli may control behavior at the expense of control by other equally relevant, concurrently available stimuli (*stimulus overselectivity*;[1] Lovaas et al., 1971). For example, in social interactions, an individual may respond only to the words spoken, despite the presence of other nonverbal cues (e.g., body language) that provide information about how to respond appropriately. In a seminal demonstration of such overselectivity, Lovaas et al. (1971) showed that, after learning to respond to a compound stimulus composed of visual (red light), auditory (white noise), and tactile (blood pressure cuff) elements, children with autism spectrum disorder (ASD) only responded to one element, whereas children with general learning disabilities and typically developing children responded to two or three elements, respectively (see also Frankel et al., 1984; Koegel & Schreibman, 1977; Koegel & Wilhelm, 1973; Kolko et al., 1980; Leader et al., 2009; Matthews et al., 2001; Reed, 2012; Schreibman & Lovaas, 1973). Similarly, Rincover and Koegel (1975) found that after children with ASD learned to follow simple instructions, instruction-following failed to generalize to different contexts (e.g., different rooms or therapists) for some children because their behavior was controlled by a restricted set of stimuli. Thus, stimulus overselectivity can lead to slow or restricted learning that does not maintain over time or generalize across contexts (see Ploog, 2010, for a review).

[1] The term 'overselectivity' implies that such selective stimulus control is necessarily or always maladaptive. This may not be so. For example, driving behavior generally comes under control of the color of traffic lights and not their vertical location, and this is efficient rather than problematic. Nevertheless, the term "overselectivity" rather than simply "selectivity" has become the norm in this literature, and we will adopt it here.

These effects of overselectivity are not limited to ASD; growing evidence shows that overselectivity also occurs in individuals with Williams syndrome (Bernardino et al., 2012; Porter & Coltheart, 2006), Fragile X syndrome (Ballantyne & Núñez, 2016), and traumatic brain injury (McHugh & Wood, 2013; Wayland & Taplin, 1985). In nonclinical populations, overselectivity has been shown to occur in preschool-aged children (Kelly & Reed, 2021; S. Reed et al., 2013), elderly adults (Kelly et al., 2016; McHugh & Reed, 2007; McHugh et al., 2010), and under conditions of high cognitive load in healthy young adults (Broomfield et al., 2010; Gomes-Ng et al., 2023; Reed & Gibson, 2005; Reed, Reynolds, & Fermandel, 2012; Reynolds & Reed, 2018). Therefore, research examining stimulus overselectivity is relevant to understanding the factors governing stimulus control across a range of populations and contexts.

Why does stimulus overselectivity occur? One perspective suggests that individuals do not attend to all stimuli during initial learning, and so only stimuli that are attended to control behavior (Dube, 2009; Dube et al., 1999, 2003, 2010; Lovaas & Schreibman, 1971; Lovaas et al., 1971; Reed, 2019). In support of this, some studies have found positive correlations between eye-gaze duration and stimulus control (stimuli observed for longer exert stronger control) in individuals with intellectual disabilities (e.g., Dube et al., 1999, 2010). Also, attention-based interventions, such as differentially reinforcing an observing response to each stimulus, may reduce overselectivity, although these intervention effects appear to fade over time (Dube & McIlvane, 1999; Dube et al., 2003, 2010; Farber et al., 2017; Reed, Altweck, et al., 2012; Walpole et al., 2007). Thus, remediating attention deficits appears to reduce overselectivity only in the short term, suggesting that other mechanisms may also underlie overselectivity (see Reed, Savile, & Truzoli, 2012, for a similar suggestion).

Alternatively, the *comparator* hypothesis asserts that individuals attend to and learn about all stimuli, but learning is not always expressed behaviorally due to competition for control between stimuli. As a result, some stimuli are selected (via a "comparator" mechanism) to exert stronger control (Reed, 2011; see also Escobar et al., 2024; Miller & Matzel, 1988). Therefore, whereas the attention-deficit view implies that overselectivity reflects the complete absence of control by some stimuli (because they were never attended to nor learned about in the first place), the comparator hypothesis suggests that overselectivity arises because control by some stimuli is masked by other stimuli. In other words, the attention-deficit mechanism operates at the stimulus *input* stage (i.e., what is initially learned), whereas the comparator mechanism operates at the behavioral *output* stage (i.e., what is expressed behaviorally). These theories make different predictions about the conditions under which control by stimuli exerting little to no control (hereafter, *underselected* stimuli) will emerge. According to attention-deficit theory, underselected stimuli can only control behavior after interventions that directly increase attention to those stimuli, whereas comparator theory predicts that control by underselected stimuli can emerge—*without* further training with those stimuli—under certain conditions.

Converging evidence from the associative-learning and operant literatures demonstrates such emergence of control by underselected stimuli after extinction of an overselected stimulus. For example, after a compound stimulus AX is paired with an unconditioned stimulus (e.g., food), extinction of the stimulus element supporting a stronger conditioned response (e.g., A) results in an increase in responding to the other element (e.g., X) despite no further training with the latter (*retrospective revaluation*; Kaufman & Bolles, 1981; Matzel et al., 1985; Shanks, 1985; see Miller & Witnauer, 2016, for a review). Similarly, Broomfield et al. (2010) found that in a simultaneous discrimination task, extinguishing choice of an overselected element increased

control over choice by an underselected element. In their study, participants learned to choose an S+ compound stimulus (e.g., AB) over an S- compound (e.g., CD), and then they chose between individual stimulus elements (e.g., B vs. D) in an unreinforced testing phase. Choice of the overselected element—defined as the S+ element, A or B, that was chosen more frequently during testing—was then extinguished and choice of a novel element was reinforced in a 'revaluation' phase. Compared with pre-revaluation testing, choice of the overselected element was lower and choice of the underselected element was higher in post-revaluation retesting (see also Gomes-Ng et al., 2023; Kelly et al., 2015; Leader et al., 2009; Reed et al., 2009; Reed, Reynolds, & Fermandel, 2012; Reynolds & Reed, 2018). Such post-revaluation changes in stimulus control are thought to reflect the change in contingency (i.e., extinction of the overselected element) during revaluation. In support of this, changes in choice are generally not apparent for a control compound whose elements do not undergo revaluation (e.g., Gomes-Ng et al., 2023; Reed et al., 2009; Reed, Reynolds, & Fermandel, 2012) or in a control group for whom choice of the overselected element continues to be reinforced during revaluation (Gomes-Ng et al., 2023).

One question that currently remains unanswered is whether it is extinction specifically, or simply the reduction in reinforcer rate associated with the overselected element (a 'worsening' of reinforcement conditions;[2] Lattal et al., 2017), that causes changes in stimulus control. In other words, would similar changes occur if reinforcers for the overselected element were reduced, but not completely removed, during revaluation? This is an important question given the applied relevance of research on overselectivity, as complete extinction may not always be possible in

[2] Henceforth, for brevity, we will refer to the extent of reduction in reinforcer rate using the term "degree of extinction." A greater degree of extinction implies a larger reduction in reinforcer rate associated with the overselected stimulus (i.e., a reinforcer rate closer to zero during the revaluation phase).

applied settings (e.g., Trump et al., 2020; Vollmer & Athens, 2011; Vollmer et al., 2020). On one hand, findings showing that compound-stimulus dimensions associated with higher reinforcer rates exert stronger control than other dimensions (e.g., Davison & Elliffe, 2010; Shahan & Podlesnik, 2006) suggest that larger decreases in reinforcer probabilities for the overselected element will produce concomitant larger increases in control by the underselected element. On the other hand, Reed, Reynolds, and Fermandel (2012) showed that increasing the number of extinction trials during revaluation weakened control by the overselected element but had little effect on the emergence of control by the underselected element. This suggests that the extent of extinction during revaluation may have graded effects on control by the overselected, but not underselected, element.

It is currently unclear how reducing, but not completely eliminating, reinforcers for the overselected element during revaluation will affect stimulus control. Therefore, the aim of the present experiment was to investigate how the degree of extinction during revaluation affects post-revaluation changes in stimulus control. We arranged a typical revaluation procedure, in which participants learned to choose a compound S+ stimulus over a compound S- stimulus, and then the overselected S+ element underwent revaluation. Unlike previous studies, in which choice of the overselected element was always extinguished completely during revaluation, the overselected element was reinforced with a probability ranging from 0 (i.e., complete extinction) to 1 (i.e., no extinction at all) during the revaluation phase in the present experiment. To assess how the degree of extinction affected stimulus control, we compared choice of each element in pre-revaluation testing with post-revaluation retesting. The attention-deficit view predicts little to no change in choice of the non-revalued elements, whereas comparator theories predict that choice of the non-revalued elements should change systematically. Furthermore, if control by the

underselected element only changes when the overselected element is completely extinguished, this would suggest that the complete removal of reinforcers is responsible for post-revaluation changes in stimulus control. In contrast, if changes in control by the underselected element are graded according to the degree of extinction, this would suggest that the *reduction* (rather than elimination) of reinforcers is sufficient to produce post-revaluation changes in stimulus control.

METHOD

Participants

We recruited 459 students enrolled at a large university in New Zealand and without a diagnosis of ASD to participate in exchange for course credit. Participants were recruited via the university's SONA system, where the study was advertised to students enrolled in psychology or communication courses. The mean age of 412 of the participants was 20.32 years (*SD* = 5.61); the remaining 47 participants declined to provide age data. Gender was not recorded.

We did not collect data on preexisting clinical conditions from participants. However, because some findings suggest that the mechanisms underlying stimulus overselectivity may differ between individuals with and without ASD (see Kelly et al., 2015; Reed, 2011, 2019; Reed & Gibson, 2005), all participants completed the Autism Quotient questionnaire (AQ; Baron-Cohen et al., 2001). The mean AQ score was 19.58 (*SD* = 7.47). Participants with AQ scores above 32 (which suggests clinically significant symptoms of ASD) were excluded from data analysis.

The program running the experiment randomly split participants into five groups (*N* = 97, 82, 93, 94, and 93 in Groups 1 to 5, respectively).

Materials and procedure

All procedures were approved by The University of Auckland Human Participants Ethics Committee (ref. 022074).

The experiment ran on Psytoolkit (Stoet, 2010, 2017), an online platform for running psychological research. Participants were sent a link to the study, and they completed one experimental session remotely on their own device. After providing informed consent, participants read the following instructions before beginning the session:

> *"In this experiment, you will make choices between different symbols. The symbols will appear in rectangles on the left and right hand sides of the screen. Choose by clicking with your mouse or touching with your finger (if on a touchscreen device).*
>
> *Sometimes, the program will tell you if you chose correctly. Other times, there will be no feedback. Your goal is to make correct choices. The more correct choices you make, the shorter the experiment will be."* (The latter instruction reflected the performance criterion in the training phase and was also included to help motivate participants to learn the discrimination.)

Experimental sessions consisted of five phases (practice, training, testing, revaluation, and retesting), in which participants chose between briefly presented S+ and S- stimuli. Figure 1A depicts a typical trial. Each trial began with a 500-ms intertrial interval, after which a fixation cross appeared in the center of the screen for 500 ms. Then, the S+ and S- stimuli were presented inside rectangles on the left and right sides of the screen. The stimuli were presented for 150 ms, after which they disappeared and the prompt "Respond now" appeared at the bottom of the screen. Participants chose a stimulus by clicking its rectangle. Feedback was provided with an

overall probability of .9 (i.e., in about 90% of trials) in the practice, training, and revaluation phases. If feedback was arranged in a trial, correct choices were followed by positive feedback (i.e., a *reinforcer*; "Correct" at the bottom of the screen for 1 s) and incorrect choices were followed by negative feedback ("Wrong!"). In the testing and retesting phases, no feedback was provided.

In all phases, the locations of the S+ and S- stimuli were randomized across trials.

Practice phase

To familiarize participants with the procedure, they first completed a practice phase in which they chose between an apple (the correct, S+, stimulus) and a balloon (the S- stimulus). After 20 practice trials, the experiment proper began.

Training phase

In the training phase, participants chose between compound stimuli. There were two pairs of compound stimuli, AB (S+) vs. CD (S-), and EF (S+) vs. GH (S-). One of these pairs was selected randomly in each trial. Each compound stimulus was composed of two abstract symbols, which were selected randomly for each participant from a pool of 20 abstract symbols (Figure 1B). Training continued for a minimum of 40 trials (20 trials per compound-stimulus pair), and then until participants reached 90% correct for each compound-stimulus pair or 100 trials maximum, whichever occurred first.

Testing phase

After training, participants completed a testing phase in which they chose between individual S+ and S- elements. There were eight pairs of elements (A vs. C, A vs. D, B vs. C, B vs. D, E vs. G,

E vs. H, F vs. G, and F vs. H), and each pair was presented five times, for a total of 40 test trials. Trials were presented in random order. No feedback was provided.

Revaluation phase

The program chose an overselected element for revaluation by randomly selecting one S+ compound (i.e., AB or EF) and then identifying the element of that compound that was chosen more frequently during the testing phase (i.e., the overselected element of that compound). For example, if compound AB was selected, then the element (A or B) that was chosen more frequently during testing was selected for revaluation. If both elements were chosen equally often, the program evaluated the other compound stimulus and used the most-selected element of that compound for revaluation. If the elements of the other compound were also chosen equally often during testing (i.e., the participant displayed no overselectivity at all), the program randomly selected any element from one of the S+ compounds for revaluation. This process ensured that whenever possible, the element selected for revaluation was indeed an overselected element (see Gomes-Ng et al., 2023). Hereafter, the compound stimulus whose overselected element underwent revaluation is termed the *revalued* compound, and the other compound stimulus is termed the *control* compound.

In the revaluation phase, participants were presented with a choice between the overselected element that was selected for revaluation and one of four novel elements. The novel elements were randomly selected from the pool of symbols (Figure 1B). Thus, there were four possible pairs of stimuli (the overselected element vs. one of the four novel elements), and one pair was selected randomly in each trial. The reinforcer ratio for choosing the overselected and novel elements varied between groups of participants. Specifically, the ratio of overselected:novel

reinforcers was 0:1 (complete extinction of the overselected element), 2:7, 1:1, 7:2, and 1:0 (no reinforcers for the novel element) in Groups 1 to 5, respectively.

Reinforcers were arranged independently. In each trial, the program first arranged feedback with a probability of .9 (i.e., as in the training phase). If feedback was arranged in a trial, a reinforcer was then allocated to the overselected or novel element according to the ratio arranged for that participant group. For example, if feedback was arranged for Group 2, then the reinforcer was allocated to the overselected element with a probability of 2/9 (about .22) and to the novel element with the complementary probability (7/9). If the reinforcer was allocated to the overselected element, then choice of that element resulted in positive feedback and choice of the novel element resulted in negative feedback. The reverse occurred if the reinforcer was allocated to the novel element.

The revaluation phase continued for 72 trials, split equally between the four pairs of stimuli.

Retesting phase

After revaluation, participants underwent a second testing phase, which was identical to the first testing phase.

Data analysis

Data are available for download in the supplementary materials.

Exclusions

Following previous studies that have investigated overselectivity in nonclinical populations (e.g., Broomfield et al., 2010; Gomes-Ng et al., 2023; Reynolds & Reed, 2018), we excluded participants with AQ scores above 32, as such scores suggest clinically significant symptoms of autism. This resulted in the removal of 33 participants' data.

There were 92 participants who did not reach the performance criterion of 90% correct within 100 trials during training. Of these, 76 participants did appear to learn the discrimination, responding correctly in at least 75% of training trials (and 66 out of the 76 participants reached at least 80% correct). The other 16 participants who did not reach the performance criterion responded correctly in less than 75% (M = 66.73%, SD = 0.08) of training trials. We chose to exclude these 16 participants because we required evidence of control by the compound stimuli in order to assess the effects of revaluation on stimulus control.

After these exclusions, there were 410 participants in the final sample (N = 81, 76, 86, 83, and 84 in Groups 1 to 5, respectively).

Training performance

We calculated the number of trials and percentage choice of the correct (S+) compound in the training phase and compared these measures across groups using a one-way ANOVA.

Overselectivity levels

We analyzed the frequency of choice of each S+ element (A, B, E, and F) during testing, and then categorized elements based on whether they were the most selected or least selected element of each compound. For example, if element A was chosen more often than element B during testing, then A was labeled the *over*selected element and B was labeled the *under*selected element for compound AB. We calculated overselectivity levels for each compound by taking the difference in percentage choice between the overselected and underselected elements. Thus, an overselectivity level of 0 reflects no difference in choice of the elements (i.e., no overselectivity), and larger values reflect greater overselectivity. We compared overselectivity levels between compounds and groups using a bootstrapped linear mixed model.

To determine whether the level of overselectivity for each compound differed from the level expected based on chance alone, we compared obtained overselectivity levels with expected levels, which were calculated based on the binomial distribution (Gomes-Ng et al., 2023; Reed, Reynolds, & Fermandel, 2012; Reynolds & Reed, 2011). For each compound, we calculated the mean probability, p, of choosing the overselected and underselected elements during testing, and then used p to calculate the probability of choosing the overselected element x times in 10 trials and the underselected element y times in 10 trials, $P(X \cap Y)$, where $X \sim Binomial(x \in [0, 10], p)$ and $Y \sim Binomial(y \in [0, 10], p)$. These binomial probabilities were entered into an 11 x 11 contingency table, and the contents of this table were multiplied by another contingency table containing the absolute $|x - y|$ difference for each combination of x and y. Summing all cells of the resulting table and then multiplying by 10 produced the expected difference between percentage choice of the overselected and underselected elements. We then compared obtained and expected levels using Wilcoxon signed-rank tests (a nonparametric alternative to the paired-samples t-test).

Effects of revaluation on choice

We aggregated the number of obtained reinforcers (i.e., positive feedback) for choosing the novel and overselected elements during revaluation and used these to calculate the obtained proportion of reinforcers for choosing the overselected element during revaluation. To assess the effects of revaluation on stimulus control, we calculated percentage choice of the overselected element during the revaluation phase, as well as the percentage *change* in choice from testing to retesting (i.e., percentage choice of each S+ element after revaluation minus percentage choice before revaluation).

One challenge with assessing the effects of revaluation on choice is that any such changes may reflect regression to the mean effects, especially for participants with more extreme pre-revaluation choice values (i.e., percentage choice closer to 0% or 100%). Regression to the mean predicts that such choice will become less extreme from testing to retesting. Moreover, levels of pre-revaluation choice effectively determine the amount of change that can occur after revaluation due to floor and ceiling effects (i.e., choice can never decrease below 0% or increase above 100%). Thus, to control for pre-revaluation choice, we ran bootstrapped linear regressions with post-revaluation choice of the overselected or underselected element as the dependent variable, the obtained proportion of reinforcers for the overselected element as a predictor, and pre-revaluation choice as a covariate. We also added overselectivity level (i.e., the difference in choice proportion between the overselected and underselected element) as a covariate because previous research suggests that overselectivity levels influence post-revaluation changes in choice (Broomfield et al., 2010; Gomes-Ng et al., 2023). This allowed us to assess the effects of changes in contingency on post-revaluation choice, independently of any effects of pre-revaluation choice and overselectivity levels.

Statistical analyses

All statistical analyses were conducted using JASP (JASP Team, 2018). Where relevant, we tested the assumptions of parametric statistical tests (e.g., normality) before running analyses. Such tests (e.g., Shapiro-Wilk, Levene's test for equality of variances) indicated that normality was violated for all measures except the number of training trials and training accuracy. Thus, we used nonparametric tests for all statistical analyses except those involving training data. For all bootstrapped tests, we used 5,000 samples.

RESULTS

Training trials to reach criterion

The mean number of training trials in Groups 1 to 5 was 71.15 (*SEM* = 3.44), 59.46 (*SEM* = 3.57), 65.87 (*SEM* = 3.21), 58.12 (*SEM* = 3.42), and 65.33 (*SEM* = 3.31), respectively. A one-way ANOVA comparing number of training trials across groups just reached statistical significance, $F(4, 405) = 2.40$, $p = .049$, $\eta_p^2 = 0.02$. Mean percentage correct in Groups 1 to 5 was 91.10% (*SEM* = 0.51), 91.75% (*SEM* = 0.46), 91.85% (*SEM* = 0.48), 92.36% (*SEM* = 0.51), and 91.33% (*SEM* = 0.43), respectively. There were no significant differences in percentage correct during training across groups, $F(4, 405) = 1.06$, $p = .379$, $\eta_p^2 = 0.01$. Thus, overall, training performance was similar across all participant groups.

Overselectivity levels before revaluation

Figure 2 shows mean percentage choice of each element of the revalued and control compounds before and after the revaluation phase in each group. Before revaluation, overselectivity levels were greater for the revalued compound—as evidenced by the larger differences in percentage choice between the overselected and underselected elements—than for the control compound. This was because of the process we used to select a compound for revaluation, which ensured that whenever possible, the revalued compound was one for which overselectivity was apparent. A bootstrapped linear mixed model (5,000 replicates) with compound as a within-subjects fixed factor, group as a between-groups fixed factor, and participant as a random factor indicated a significant main effect of compound, $F(1, 405) = 47.02$ [95% CI 24.84, 80.59], $p < .001$, $\eta_p^2 = 0.10$, but no effect of group, $F(4, 769) = 0.66$ [95% CI 0.36, 4.74], $p = .617$, $\eta_p^2 < 0.01$, or of the compound x group interaction, $F(4, 405) = 1.95$ [95% CI 0.62, 6.90], $p = .102$, $\eta_p^2 = 0.02$. Thus,

overselectivity levels differed between compounds, but not between groups arranging different reinforcer probabilities during revaluation.

To determine whether levels of overselectivity differed from levels expected based on chance alone, we compared obtained and expected overselectivity levels using one-tailed Wilcoxon signed-rank tests. Table 1 shows mean obtained and expected levels in each group, as well as the results of the Wilcoxon signed-rank tests. Obtained levels of overselectivity were significantly larger than expected levels for both compounds and in all groups. Thus, our procedure successfully induced overselectivity.

Choice during revaluation

In the revaluation phase, choice of the overselected element of the revalued compound was followed by positive feedback (the reinforcer) with varying probabilities across groups. Figure 3A shows box-and-whisker plots of obtained reinforcer probabilities during revaluation for each element (calculated by dividing the number of reinforcers obtained for choosing an element by the number of times that element was chosen), and Figure 3B shows choice of the overselected element during the revaluation phase. Overall, median reinforcer probabilities approximated arranged probabilities in all groups, and mean percentage choice of the overselected element approximately matched the reinforcer probability. Thus, choice during the revaluation phase was controlled by the reinforcer contingencies during that phase.

However, the whiskers in Figure 3A also show that obtained reinforcer probabilities deviated from arranged probabilities for some participants. Such deviations were particularly apparent when reinforcer probabilities were less extreme (i.e., Groups 2 to 4). These deviations between arranged and obtained reinforcer probabilities occurred because we arranged reinforcers

independently, and because of the dynamical relationship between behavior and reinforcers. That is, to experience the arranged reinforcer contingencies during revaluation, participants needed to choose the overselected and novel elements and receive feedback following such choices. As a result, whether participants experienced the arranged probabilities depended on how frequently they chose each element (e.g., if a participant never or hardly chose one element, then the obtained proportion of reinforcers for that element may have been close to 0). Despite such deviations, obtained reinforcer probabilities still followed the same *pattern* as arranged probabilities (i.e., increasing for the overselected element and decreasing for the novel element from Groups 1 to 5). Thus, the deviations between arranged and obtained reinforcer probabilities did not prevent us from assessing the effects of the contingency change on post-revaluation stimulus control.

Effects of revaluation on stimulus control

To assess how revaluation affected control by the elements of each S+ compound, we compared choice during the retesting phase (i.e., after revaluation) with choice during the testing phase (i.e., before revaluation). Mean percentage choice before and after revaluation is shown in Figure 2, and Figure 4 shows the effects of revaluation more clearly by plotting the mean change in percentage choice from pre- to post-revaluation. Changes in choice for individual participants in each group are plotted in Appendix Figures A1 to A5. Negative values in Figure 4 and in Appendix Figures A1 to A5 reflect a decrease in choice of that element, whereas positive values reflect an increase. Changes in choice of the overselected element of the revalued compound were negative, and consistent with the reinforcer probabilities arranged during revaluation (Figure 3); thus, the effects of the revaluation phase maintained after that phase ended. For the underselected element of the revalued compound, choice appeared to decrease slightly when

choice of the overselected element of that compound was extinguished (Group 1), whereas choice increased slightly in the other groups. Changes in choice of the elements of the control compound tended to follow a similar pattern to the revalued compound, except that such changes were smaller for the control compound. These patterns were evident for most participants within a group, although there was variability in the magnitude, and sometimes direction, of changes in choice between participants (Appendix Figures A1 to A5).

Predictors of post-revaluation changes in stimulus control

Figure 4 appears to suggest that revaluation had small effects on choice of the underselected element of the revalued compound. Additionally, excluding Group 1, there was little difference between groups, although increases in choice of the underselected element did appear to be slightly larger when the reinforcer probability for the overselected element was higher. However, given the substantial variability in obtained reinforcer probabilities within a group (Figure 2), it is possible that aggregating data across all participants within a group masked any differential effects of obtained reinforcer probabilities on stimulus control. For example, consider Group 2 in which the arranged reinforcer probability for the overselected element was .2 but obtained probabilities ranged from 0 to .44—post-revaluation changes in choice may have differed for participants at the lower and higher ends of this range, but any such differences would not be apparent in Figure 4. Thus, to examine the effects of the change in contingency more precisely, we ran linear regressions with the obtained proportion of reinforcers for the overselected element (i.e., number of reinforcers obtained for choosing the overselected element, divided by total number of obtained reinforcers during revaluation) as a predictor of post-revaluation choice percentages.

Additionally, to account for regression to the mean effects (artifactual shifts in choice from more extreme to less extreme values over time), we added pre-revaluation choice percentages as a covariate to the linear regression model (see Barnett et al., 2005). We also included overselectivity level as a covariate, as prior research suggests that the effects of revaluation on stimulus control may differ depending on initial levels of overselectivity (Gomes-Ng et al., 2023).[3] Table 2 shows the results of the bootstrapped linear regressions (5,000 replicates). There was a significant effect of pre-revaluation choice for all four elements; thus, changes in choice after revaluation depended to an extent on choice before revaluation (a potential regression-to-the-mean effect; e.g., larger increases for participants with initially lower percentage of choice of an element). In contrast, initial overselectivity levels – the *relative* difference in percentage choice between the elements within a compound before revaluation – did not significantly predict post-revaluation choice. After controlling for pre-revaluation choice and overselectivity level, there was a significant positive relationship between obtained reinforcer proportions and post-revaluation choice of all four elements, indicating that choice increased to a slightly greater extent for participants who obtained more reinforcers for choosing the overselected element during revaluation. This relationship was strongest for the overselected element of the revalued compound, and weak for the underselected element of the revalued compound and both elements of the control compound. Thus, reinforcer probabilities during revaluation predicted, albeit

[3] We also ran the regressions with age, AQ score, number of training trials, and percentage of correct during training as additional covariates. Including these additional covariates did not alter the pattern of results: Pre-revaluation choice percentages and obtained reinforcer proportions, but not overselectivity level, remained significant predictors of post-revaluation choice percentages. None of the additional covariates were significant predictors (all ps > .05), with one exception: AQ scores predicted post-revaluation choice of the overselected element of the revalued compound (β = -.45, p = .025). AQ scores clearly had little systematic effect on post-revaluation choice, particularly for the elements for which changes in choice were emergent (i.e., the underselected element of the revalued compound, and both elements of the control compound).

weakly for most elements, post-revaluation choice percentages even after controlling for regression-to-the-mean effects.

DISCUSSION

When a compound-stimulus element exerting strong control over behavior is extinguished (*revaluation*), control by the other element(s) may change despite no further training with those elements (e.g., Broomfield et al., 2010; Gomes-Ng et al., 2023; Kelly et al., 2015; Reynolds & Reed, 2018). In the present study, we asked how the degree of extinction during revaluation affects post-revaluation changes in stimulus control. We induced overselectivity in healthy adults by rapidly presenting compound S+ and S- stimuli in a simultaneous discrimination task. During the revaluation phase, participants experienced varying degrees of extinction, ranging from complete extinction to no change in contingency, for the overselected element of one S+ compound. Choice of the overselected element of the revalued compound changed in accordance with the degree of extinction, with larger decreases when choice of that element was reinforced with a lower probability during revaluation. Choice of the underselected element of the revalued compound also decreased when the overselected element was extinguished completely (Group 1), whereas it increased slightly when the overselected element was reinforced some or all of the time during revaluation (Groups 2 to 5; Figure 4). Choice of the control-compound elements followed a similar pattern, though changes for the control compound were smaller overall. Besides the difference in the direction of changes in choice between Group 1 and the other groups, the degree of extinction had only a small effect on post-revaluation changes in stimulus control. This—along with prior work showing that the number of extinction trials had little effect on the emergence of control by the underselected element (Reed, Reynolds, & Fermandel, 2012)

—suggests that post-revaluation changes in stimulus control may only be slightly (if at all) graded according to the extent of contingency change during revaluation.

Our findings make several contributions to the literature on stimulus overselectivity. Unlike previous research, in which overselectivity was induced by arranging a concurrent-load subtraction task in in-person sessions (e.g., Broomfield et al., 2010; Gomes-Ng et al., 2023; Reed, Reynolds, & Fermandel, 2012; Reynolds & Reed, 2018), our participants completed an online session in which stimuli were presented rapidly. Our procedure successfully induced overselectivity, and to levels comparable with previous studies (compare e.g., our Figure 2 with Figure 3 in Gomes-Ng et al., 2023 or Figure 2 in Broomfield et al., 2010). Thus, our findings add to the evidence showing that overselectivity can be induced in healthy adults without a diagnosis of ASD under conditions that make discrimination learning more difficult. Additionally, this study is the first (to our knowledge) to induce overselectivity in an online experimental environment. Thus, our findings also demonstrate the utility and viability of online data-collection tools for overselectivity research.

One possible explanation for the post-revaluation changes in choice in the present experiment is that such changes were artifactual. Although we cannot completely rule this out, especially as changes in choice were relatively small (although they appear similar in magnitude to previous research, see e.g., Broomfield et al., 2008; Gomes-Ng et al., 2023; Kelly et al., 2015; Reed, Reynolds, & Fermandel, 2012; Reynolds & Reed, 2018), two aspects of our results appear to suggest that such changes were not merely an artifact. First, changes in choice were consistently larger for the revalued than for the control compound, and second, choice changed in different directions for different participant groups (specifically, in Group 1 relative to the other conditions). If changes in choice were purely artifactual, they should have been similar—or at

least not systematically different—across compounds and groups. Thus, these systematic differences imply that post-revaluation changes in choice reflected the contingencies arranged during revaluation.

Two theories of overselectivity make different predictions about the effects of revaluation on stimulus control. The attention-deficit view posits that overselectivity reflects a failure to observe or attend to some stimulus elements during training and thus changes in control by ignored elements can only occur after further training with those elements (e.g., Dube, 2009; Reed, 2019). In contrast, comparator theories state that individuals attend to and learn about all compound-stimulus elements, but competition for stimulus control results in some elements exerting stronger control than others (e.g., Reed, 2011). As a result, changing the reinforcer contingencies for the overselected element may alter control by other, underselected elements. Even though we presented stimuli rapidly (which may have made it harder to observe or attend to all stimulus elements), our findings are clearly consistent with comparator theories, as systematic post-revaluation changes in choice were apparent. Thus, despite differences in the method used to induce overselectivity, similar comparator mechanisms appear to underlie overselectivity in the present and past experiments.

Although the mechanisms underlying overselectivity appear similar, there are two interesting differences between our findings and those of past research. First, in past studies, complete extinction of the overselected element during revaluation produced larger increases in choice of the underselected element for participants with greater initial overselectivity levels (Broomfield et al., 2010; Gomes-Ng et al., 2023). In contrast, overselectivity levels did not predict post-revaluation choice in the present experiment. This may be because we controlled for pre-revaluation choice (by adding it as a covariate in the regression model), a measure which is

necessarily related to overselectivity levels (e.g., greater choice of the underselected element necessarily implies less overselectivity). However, when we reanalyzed Gomes-Ng et al.'s (2023) data with pre-revaluation choice as an added covariate, overselectivity levels still predicted changes in choice. This suggests that even when overselectivity is induced to similar levels (as in the present and past experiments), procedural differences—perhaps relating to how overselectivity is induced—may influence the extent to which overselectivity levels influence post-revaluation changes in choice. Given the lack of research in this area, it is unclear which procedural aspects may contribute. There are two major procedural differences between the present and past experiments—we used rapidly presented stimuli (vs. concurrent-load tasks) in an online (vs. in person) environment. Thus, as a starting point, future work could directly compare these procedural aspects to determine how they may influence the effects of overselectivity levels on post-revaluation changes in choice.

The second difference between our results and past findings is that choice of the underselected element decreased after complete extinction of the overselected element in the present experiment, whereas choice of the underselected element *increased* under the same contingencies in previous studies (e.g., Broomfield et al., 2010; Gomes-Ng et al., 2023; Reed, Reynolds, & Fermandel, 2012; Reynolds & Reed, 2018). If similar mechanisms underlie overselectivity in these studies, why did extinction of the overselected element have opposite effects on choice of the underselected element? Some evidence suggests that learned relations between compound-stimulus elements (sometimes termed 'within-compound associations') moderate the effects of extinction on stimulus control (see Dickinson & Burke, 1996; Gomes-Ng et al., 2023; Miller & Witnauer, 2016; Reed & Quigley, 2019, for discussions). Specifically, retrospective revaluation appears more likely under conditions that engender weaker within-

compound relations, such as when compound stimuli are presented for a short duration (e.g., 2 s) during training or stimulus elements are more physically or functionally disparate. In contrast, under conditions that promote stronger within-compound relations (e.g., longer stimulus presentations, greater similarity between elements), control by the underselected element decreases (*mediated extinction*; e.g., Liljeholm & Balleine, 2009; Reynolds & Reed, 2018; Sissons et al., 2009). Compared with past research, our participants may have formed stronger within-compound relations, perhaps because the discrimination was easier to acquire due to the absence of a concurrent-load task. This may explain why choice of the underselected element decreased when the overselected element was completely extinguished (Group 1) and increased when the overselected element was always reinforced (Group 5; a mediated 'acquisition' effect akin to mediated extinction). Similarly, participants may have formed strong relations between the revalued and control compounds because both were trained in the same context and were physically (i.e., in appearance; Figure 1B) and functionally (i.e., both were S+ stimuli) similar (see e.g., Minster et al., 2006; Schenk, 1994). As a result, the effects of revaluation may have generalized to the control compound via *between*-compound relations (see Gomes-Ng et al., 2023, for a similar discussion). To test this, future work could replicate the current experiment and vary the stimulus duration or physical disparity of the stimulus elements. If shorter durations or more disparate elements result in retrospective revaluation (i.e., choice for the underselected element increases after complete extinction of the overselected element), and little change in choice of the control-compound elements, this would lend support to our suggestion that within- and between-compound relations contributed to the present results.

However, this line of reasoning is inconsistent with the results of groups arranging partial extinction (i.e., Groups 2 to 4), because choice of the overselected and underselected elements

changed in opposite directions in these groups (i.e., the typical retrospective-revaluation effect; Figure 4). Perhaps different mechanisms operate when the contingencies during revaluation are certain (choice of the overselected element is never or always reinforced) versus uncertain (choice of the overselected element is sometimes reinforced; see e.g., Carroll et al., 2010, for a brief discussion), although *what* those mechanisms remains unclear. What is evident from our findings is that the effects of extinguishing an overselected element on control by underselected elements are not as straightforward as they may first appear. Therefore, future research should establish the replicability and generality of our findings, especially in the typical overselectivity procedure (i.e., varying reinforcer probabilities during revaluation in an in-person session with a concurrent-load task). This would provide insight into whether our findings relate to a specific aspect of our procedure, or whether they reflect more general effects of complete (Group 1), partial (Groups 2 to 4), or no (Group 5) extinction of the overselected element.

Our finding of retrospective revaluation when the overselected element was partially extinguished (Figure 4) is of applied relevance. Stimulus overselectivity can hamper the success of behavioral interventions, particularly if irrelevant stimuli, or a restricted subset of relevant stimuli, exert near-exclusive control (Ploog, 2010). Under such conditions, remediating overselectivity is important to ensure that target behaviors are learned, maintained, and generalized. While previous research suggests that complete extinction of an overselected stimulus enhances control by underselected stimuli (e.g., Broomfield et al., 2010; Gomes-Ng et al., 2023; Kelly et al., 2015; Leader et al., 2009; Reed et al., 2009; Reed, Reynolds, & Fermandel, 2012; Reynolds & Reed, 2018), our results show that such post-revaluation changes in stimulus control may be achieved without complete extinction. This is important because complete extinction may not always be practical or possible (e.g., if sensory reinforcers are

involved; see Trump et al., 2020; Vollmer & Athens, 2011; Vollmer et al., 2020, for discussions), nor may complete extinction be ideal if some control by the overselected stimulus is still desired. Our results tentatively suggest that simply reducing the reinforcer rate for responding to an overselected stimulus may weaken (but not completely eliminate) control by the overselected stimulus and enhance control by underselected stimuli, resulting in a more equal division of stimulus control (Figure 2). To be sure, replicating our study with clinical populations displaying overselectivity (e.g., ASD) is critical to establish whether partial extinction can indeed remediate overselectivity in clinical settings.

Our findings highlight the need for better control conditions against which to compare the effects of revaluation. The typical approach is to include a control compound that does not undergo revaluation, and to assume that changes in choice of that compound are purely artifactual (e.g., Broomfield et al., 2008, 2010; Gomes-Ng et al., 2023; Reed et al., 2009). Here, we found that choice of the control-compound elements changed similarly to the revalued-compound elements (Figure 4). Although we cannot rule out an artifactual account, the systematic differences in choice of the control-compound elements between groups do appear to suggest that such changes were not solely artifactual but instead reflected *generalization* of the effects of revaluation. This suggests that a control compound alone may not be the best way to assess the effects of revaluation, particularly under conditions that are likely to promote strong within- or between-compound relations. Instead, future work could include a control condition in which participants choose between novel elements during revaluation; in this condition, any post-revaluation changes in choice would be entirely artifactual, because no additional training took place with the revalued compound. Furthermore, given that some research suggests that introducing novel stimuli can alter control by familiar stimuli (a ‘novelty effect’; e.g., Sutherland, 1961; Warran &

McGonigle, 1969), future studies should investigate the extent to which such novelty effects contribute to post-revaluation changes in stimulus control. For instance, to control for novelty, participants could be exposed to the stimuli to be introduced during revaluation as part of a prior or concurrent task.

In conclusion, our results demonstrate that extinguishing an overselected element can alter control by both the overselected and underselected elements. This adds to the evidence that overselectivity reflects a comparator mechanism, rather than an attention deficit. Importantly, our findings tentatively suggest that complete extinction of the overselected element is not the only way to increase control by underselected elements. Instead, it may be possible to increase such control by simply reducing reinforcers for the overselected element. This has important applied implications, given that complete extinction is not always practical, possible, or desirable in natural environments. Our findings point to several important avenues for future research, including arranging stronger control conditions, testing the replicability and generality of our results, and examining whether partial extinction can remediate overselectivity in clinical populations. Exploring these avenues will advance our understanding of why and how contingency changes for responding to an overselected stimulus affect stimulus control.

AUTHOR CONTRIBUTIONS

Gomes-Ng: Conceptualization, methodology, software, investigation, formal analysis, visualization, supervision, writing – original draft, writing – review and editing, funding acquisition. **Cowie:** Writing – review and editing. **Elliffe:** Methodology, writing – review and editing.

ACKNOWLEDGMENTS

The authors thank Georgia Keys for their help in collecting data and Dr. John Bai for his helpful comments on a draft of the manuscript. This research was supported by a Lottery Health Research Postdoctoral Fellowship awarded to Stephanie Gomes-Ng (LHR-2022-185377).

CONFLICT OF INTEREST STATEMENT

The authors declare not conflicts of interest.

DATA AVAILABILITY STATEMENT

Data are available for download in the supplementary materials.

ETHICS APPROVAL

This research was approved by The University of Auckland Participants Ethics Committee (ref. 022074).

REFERENCES

Ballantyne, C. J., & Núñez, M. (2016). Developmental trajectories of hierarchical visuo-spatial processing in Fragile X syndrome and ASD: Within- and cross-syndrome variability. *Research in Developmental Disabilities*, *51-52*, 103–115. https://doi.org/10.1016/j.ridd.2015.12.016

Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology*, *34*(1), 215–220. https://doi.org/10.1093/ije/dyh299

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*, 5–17. https://doi.org/10.1023/A:1005653411471

Bernardino, I., Mouga, S., Almeida, J., van Asselen, M., Oliveira, G., & Castelo-Branco, M. (2012). A direct comparison of local-global integration in autism and other developmental disorders: Implications for the central coherence hypothesis. *PloS ONE*, *7*(6), e39351. https://doi.org/10.1371/journal.pone.0039351

Broomfield, L., McHugh, L., & Reed, P. (2008). Re-emergence of under-selected stimuli, after the extinction of over-selected stimuli in an automated match to samples procedure. *Research in Developmental disabilities*, *29*(6), 503–512. https://doi.org/10.1016/j.ridd.2007.09.001

Broomfield, L., McHugh, L., & Reed, P. (2010). Factors impacting emergence of behavioral control by underselected stimuli in humans after reduction of control by overselected stimuli. *Journal of the Experimental Analysis of Behavior*, *94*(2), 125–133. https://doi.org/10.1901/jeab.2010.94-125

Carroll, C., Cheng, P., & Lu, H. (2010). Uncertainty in causal inference: The case of retrospective revaluation. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *32*, 1076–1081.

Davison, M., & Elliffe, D. (2010). Divided stimulus control: A replication and a quantitative model. *Journal of the Experimental Analysis of Behavior*, *94*(1), 13–23. https://doi.org/10.1901/jeab.2010.94-13

Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *Quarterly Journal of Experimental Psychology*, *49*(1), 60–80. https://doi.org/10.1080/713932614

Dube, W. V. (2009). Stimulus overselectivity in discrimination learning. In P. Reed (Ed.), *Behavioral theories and interventions for autism* (pp. 23–46). Nova Science Publishers.

Dube, W. V., Dickson, C. A., Balsamo, L. M., O'Donnell, K. L., Tomanari, G. Y., Farren, K. M., Wheeler, E. E., & McIlvane, W. J. (2010). Observing behavior and atypically restricted stimulus control. *Journal of the Experimental Analysis of Behavior*, *94*(3), 197–313. https://doi.org/10.1901/jeab.2010.94-297

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*, *17*, 267–271.

Dube, W. V., Lombard, K. M., Farren, K. M., Flusser, D. S., Balsamo, L. M., Fowler, T. R., & Tomanari, G. Y. (2003). Stimulus overselectivity and observing behavior in individuals with mental retardation. In S. Soraci & K. Murata-Soraci (Eds.), *Visual information processing* (pp. 109–123). Praeger.

Dube, W. V., & McIlvane, W. J. (1999). Reduction of stimulus overselectivity with nonverbal differential observing responses. *Journal of Applied Behavior Analysis*, *32*(1), 25–33. https://doi.org/10.1901/jaba.1999.32-25

Escobar, M., Bell, Z. K., & Arcediano, F. (2024). Stimulus selection and stimulus competition. In H. S. Roane, A. R. Craig, V. Saini, & J. E. Ringdahl (Eds.), *Behavior analysis: Translational perspectives and clinical practice* (pp. 146-169). Guilford.

Farber, R. S., Dickson, C. A., & Dube, W. V. (2017). Reducing overselective stimulus control with differential observing responses. *Journal of Applied Behavior Analysis*, *50*(1), 87–105. https://doi.org/10.1002/jaba.363

Frankel, F., Simmons, J. Q., Fichter, M., & Freeman, B. J. (1984). Stimulus overselectivity in autistic and mentally retarded children—a research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *25*(1), 147–155. https://doi.org/10.1111/j.1469-7610.1984.tb01727.x

Gomes-Ng, S., Kim, P. B. C., Cowie, S., & Elliffe, D. (2023). Revaluation of overselected stimuli: Emergence of control by underselected stimuli depends on degree of overselectivity. *Journal of the Experimental Analysis of Behavior*, *120*(2), 155–170. https://doi.org/10.1002/jeab.850

JASP Team. (2018). *JASP* (Version 0.19.2) [Computer software]. https://jasp-stats.org

Kaufman, M. A., & Bolles, R. C. (1981). A nonassociative aspect of overshadowing. *Bulletin of the Psychonomic Society*, *18*, 318–320. https://doi.org/10.3758/BF03333639

Kelly, M. P., Leader, G., & Reed, P. (2015). Stimulus over-selectivity and extinction-induced recovery of performance as a product of intellectual impairment and autism severity. *Journal of Autism and Developmental Disorders*, *45*, 3098–3106. https://doi.org/10.1007/s10803-015-2466-x

Kelly, M. P., Leader, G., & Reed, P. (2016). Factors producing over-selectivity in older individuals. *AGE*, *38*, 63. https://doi.org/10.1007/s11357-016-9926-x

Kelly, M. P., & Reed, P. (2021). Age effects on the development of stimulus over-selectivity are mediated by cognitive flexibility and selective attention. *International Journal of Behavioral Development*, *45*(1), 89–96. https://doi.org/10.1177/0165025420949702

Koegel, R. L., & Schreibman, L. (1977). Teaching autistic children to respond to simultaneous multiple cues. *Journal of Experimental Child Psychology*, *24*(2), 299–311. https://doi.org/10.1016/0022-0965(77)90008-X

Koegel, R. L., & Wilhelm, H. (1973). Selective responding to the components of multiple visual cues by autistic children. *Journal of Experimental Child Psychology*, *15*(3), 442–453. https://doi.org/10.1016/0022-0965(73)90094-5

Kolko, D. J., Anderson, L., & Campbell, M. (1980). Sensory preference and overselective responding in autistic children. *Journal of Autism and Developmental Disorders*, *10*, 259–271. https://doi.org/10.1007/BF02408285

Leader, G., Loughnane, A., McMorland, C., & Reed, P. (2008). The effect of stimulus salience on over-selectivity. *Journal of Autism and Developmental Disorders*, *39*, 330–338. https://doi.org/10.1007/s10803-008-0626-y

Liljeholm, M., & Balleine, B. W. (2009). Mediated conditioning versus retrospective revaluation in humans: The influence of physical and functional similarity of cues. *Quarterly Journal of Experimental Psychology*, *62*(3), 470–482. https://doi.org/10.1080/17470210802008805

Lovaas, O. I., & Schreibman, L. (1971). Stimulus overselectivity of autistic children in a two stimulus situation. *Behaviour Research and Therapy*, *9*(4), 305–310. https://doi.org/10.1016/0005-7967(71)90042-8

Lovaas, O. I., Schreibman, L., Koegel, R., & Rehm, R. (1971). Selective responding by autistic children to multiple sensory input. *Journal of Abnormal Psychology*, *77*(3), 211–222. https://doi.org/10.1037/h0031015

Matthews, B., Shute, R., & Rees, R. (2001). An analysis of stimulus overselectivity in adults with autism. *Journal of Intellectual & Developmental Disability*, *26*(2), 161–176. https://doi.org/10.1080/13668250020054477

Matzel, L. D., Schachtman, T. R., & Miller, R. R. (1985). Recovery of an overshadowed association achieved by extinction of the overshadowing stimulus. *Learning and Motivation*, *16*(4), 398–412. https://doi.org/10.1016/0023-9690(85)90023-2

McHugh, L., & Reed, P. (2007). Age trends in stimulus overselectivity. *Journal of the Experimental Analysis of Behavior*, *88*(3), 369–380. https://doi.org/10.1901/jeab.2007.88-369

McHugh, L., Simpson, A., & Reed, P. (2010). Mindfulness as a potential intervention for stimulus over-selectivity in older adults. *Research in Developmental Disabilities*, *31*(1), 178–184. https://doi.org/10.1016/j.ridd.2009.08.009

McHugh, L., & Wood, R. (2013). Stimulus over-selectivity in temporal brain injury: Mindfulness as a potential intervention. *Brain Injury*, *27*(13-14), 1595–1599. https://doi.org/10.3109/02699052.2013.834379

Miller, R. R., & Matzel, L. D. (1988). The comparator hypothesis: A response rule for the expression of associations. *Psychology of Learning and Motivation*, *22*, 51–92. https://doi.org/10.1016/S0079-7421(08)60038-9

Miller, R. R., & Witnauer, J. E. (2016). Retrospective revaluation: The phenomenon and its theoretical implications. *Behavioural Processes*, *123*, 15–25. https://doi.org/10.1016/j.beproc.2015.09.001

Minster, S. T., Jones, M., Elliffe, D., & Muthukumaraswamy, S. D. (2006). Stimulus equivalence: Testing Sidman's (2000) theory. *Journal of the Experimental Analysis of Behavior*, *85*(3), 371–391. https://doi.org/10.1901/jeab.2006.15-05

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. https://doi.org/10.1007/s10803-010-0990-2

Porter, M. A., & Coltheart, M. (2006). Global and local processing in Williams syndrome, autism, and Down syndrome: Perception, attention, and construction. *Developmental Neuropsychology*, *30*(3), 771–789. https://doi.org/10.1207/s15326942dn3003_1

Reed, P. (2012). Brief report: The effect of delayed matching to sample on stimulus over-selectivity. *Journal of Autism and Developmental Disorders*, *42*, 1515–1519. https://doi.org/10.1007/s10803-011-1374-y

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* (pp. 168–190). Oxford University Press.

Reed, P. (2019). Unpredictability reduces over-selective responding of individuals with ASD who have language impairments. *Research in Autism Spectrum Disorders*, *57*, 35–45. https://doi.org/10.1016/j.rasd.2018.10.006

Reed, P., Altweck, L., Broomfield, L., Simpson, A., & McHugh, L. (2012). Effect of observing-response procedures on overselectivity in individuals with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, *27*(4), 237–246. https://doi.org/10.1177/1088357612457986

Reed, P., Broomfield, L., McHugh, L., McCausland, A., & Leader, G. (2009). Extinction of over-selected stimuli causes emergence of under-selected cues in higher-functioning children with autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*, 290–298. https://doi.org/10.1007/s10803-008-0629-8

Reed, P., & Gibson, E. (2005). The effect of concurrent task load on stimulus over-selectivity. *Journal of Autism and Developmental Disorders*, *35*, 601–614. https://doi.org/10.1007/s10803-005-0004-y

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

Reed, P., Reynolds, G., & Fermandel, L. (2012). Revaluation manipulations produce emergence of underselected stimuli following simultaneous discrimination in humans. *Quarterly Journal of Experimental Psychology*, *65*(7), 1345–1360. https://doi.org/10.1080/17470218.2012.656663

Reed, P., Savile, A., & Truzoli, R. (2012). Event-related potential analysis of stimulus over-selectivity. *Research in Developmental Disabilities*, *33*(2), 655–662. https://doi.org/10.1016/j.ridd.2011.11.012

Reed, S. R., Stahmer, A. C., Suhrheinrich, J., & Schreibman, L. (2013). Stimulus overselectivity in typical development: Implications for teaching children with autism. *Journal of Autism and Developmental Disorders*, *43*, 1249–1257. https://doi.org/10.1007/s10803-012-1658-x

Rincover, A., & Koegel, R. L. (1975). Setting generality and stimulus control in autistic children. *Journal of Applied Behavior Analysis*, *8*(3), 235–246. https://doi.org/10.1901/jaba.1975.8-235

Reynolds, G., & Reed, P. (2011). Effects of schedule of reinforcement on over-selectivity. *Research in Developmental Disabilities*, *32*(6), 2489–2501. https://doi.org/10.1016/j.ridd.2011.07.011s

Reynolds, G., & Reed, P. (2018). The effect of stimulus duration on over-selectivity: Evidence for the role of within-compound associations. *Journal of Experimental Psychology: Animal Learning and Cognition*, *44*(3), 293–308. https://doi.org/10.1037/xan0000175

Schenk, J. J. (1994). Emergent relations of equivalence generated by outcome-specific consequences in conditional discrimination. *The Psychological Record*, *44*(4), 537–558. https://doi.org/10.1007/BF03395143

Schreibman, L., & Lovaas, O. I. (1973). Overselective response to social stimuli by autistic children. *Journal of Abnormal Child Psychology*, *1*, 152–168. https://doi.org/10.1007/BF00916110

Shahan, T. A., & Podlesnik, C. A. (2006). Divided attention performance and the matching law. *Learning & Behavior*, *34*, 255–261. https://doi.org/10.3758/BF03192881

Shanks, D. R. (1985). Forward and backward blocking in human contingency judgement. *Quarterly Journal of Experimental Psychology*, *37*(1b), 1–21. https://doi.org/10.1080/14640748508402082

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. https://doi.org/10.3758/LB.37.3.254

Skinner, B. F. (1953). *Science and human behavior.* Macmillan.

Stoet, G. (2010). Psytoolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*, 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). Psytoolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Sutherland, N. S. (1961). Shape discrimination by animals. *Experimental Psychology Society Monographs*, *1*, 1–68.

Trump, C. E., Ayres, K. M., Quinland, K. K., & Zabala, K. A. (2020). Differential reinforcement without extinction: A review of the literature. *Behavior Analysis: Research and Practice*, *20*(2), 94–107. https://doi.org/10.1037/bar0000169

Vollmer, T. R., & Athens, E. (2011). Developing function-based extinction procedures for problem behavior. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 317–334). Guilford.

Vollmer, T. R., Peters, K. P., Kronfli, F. R., Lloveras, L. A., & Ibañez, V. F. (2020). On the definition of differential reinforcement of alternative behavior. *Journal of Applied Behavior Analysis*, *53*(3), 1299–1303. https://doi.org/10.1002/jaba.701

Walpole, C. W., Roscoe, E. M., & Dube, W. V. (2007). Use of a differential observing response to expand restricted stimulus control. *Journal of Applied Behavior Analysis*, *40*(4), 707–712. https://doi.org/10.1901/jaba.2007.707-712

Wayland, S., & Taplin, J. E. (1985). Feature-processing deficits following brain injury: Overselectivity in recognition memory for compound stimuli. *Brain and Cognition*, *4*(3), 338–355. https://doi.org/10.1016/0278-2626(85)90026-0

TABLE 1 Expected and obtained levels of overselectivity before revaluation.

	Revalued compound			Control compound		
	Expected (*SEM*)	Obtained (*SEM*)	*W*	Expected (*SEM*)	Obtained (*SEM*)	*W*
Cond. 1	9.82 (0.08)	20.86 (0.28)	1720***	7.05 (0.09)	14.57 (0.34)	644*
Cond. 2	10.26 (0.07)	22.76 (0.24)	2231***	7.68 (0.10)	18.42 (0.35)	731***
Cond. 3	11.11 (0.07)	29.19 (0.30)	2686***	7.26 (0.08)	16.16 (0.28)	945***
Cond. 4	10.90 (0.07)	23.86 (0.23)	2491***	8.65 (0.08)	13.86 (0.23)	1136**
Cond. 5	11.14 (0.07)	28.10 (0.29)	2539***	7.51 (0.08)	13.10 (0.25)	879*

Note: *W* shows the test statistic for one-tailed Wilcoxon signed-rank tests comparing obtained and expected values. *$p < .05$; **$p < .01$; ***$p < .001$

TABLE 2 Results of bootstrapped linear regressions with post-revaluation choice percentages as the dependent variable.

	β	Bias	*SE*	*p*	95% CI
Revalued compound					
Overselected, $F(3, 406) = 63.68$, $p < .001$, $R^2 = .32$					
Rf	0.37	< .0001	0.03	< .001	[0.31, 0.43]
OS level	< .0001	< .000	0.05	.974	[-0.11, 0.10]
Prechoice	0.56	0.01	0.14	< .001	[0.25, 0.78]
Underselected, $F(3, 406) = 60.68$, $p < .001$, $R^2 = .31$					
Rf	0.11	0.00	0.03	< .001	[0.05, 0.16]
OS level	0.01	0.01	0.14	.988	[-0.30, 0.26]
Prechoice	0.60	0.01	0.13	< .001	[0.33, 0.82]
Control compound					
Overselected, $F(3, 406) = 27.98$, $p < .001$, $R^2 = .17$					
Rf	0.15	< .0001	0.03	< .001	[0.10, 0.21]
OS level	-0.10	< .0001	0.06	.061	[-0.22, 0.00]
Prechoice	0.53	0.01	0.13	< .001	[0.28, 0.78]
Underselected, $F(3, 406) = 86.53$, $p < .001$, $R^2 = .39$					
Rf	0.14	< .0001	0.03	< .001	[0.08, 0.20]
OS level	-0.16	0.01	0.14	.245	[-0.42, 0.12]
Prechoice	0.52	0.01	0.13	< .001	[0.28, 0.76]

Note: Rf = obtained reinforcer proportion for the overselected element during revaluation, OS level = level of overselectivity, Prechoice = pre-revaluation percentage choice. The *p* values and 95% confidence intervals (CI) are bias corrected.

FIGURE 1 Structure of a typical trial (A) and abstract symbols used for stimulus elements (B). Compound (two-element) stimuli were presented in training, and individual elements were presented in the testing and revaluation phases. Feedback was only provided during training and revaluation. The stimulus elements used for each participant were selected from the pool of symbols shown in Panel B.

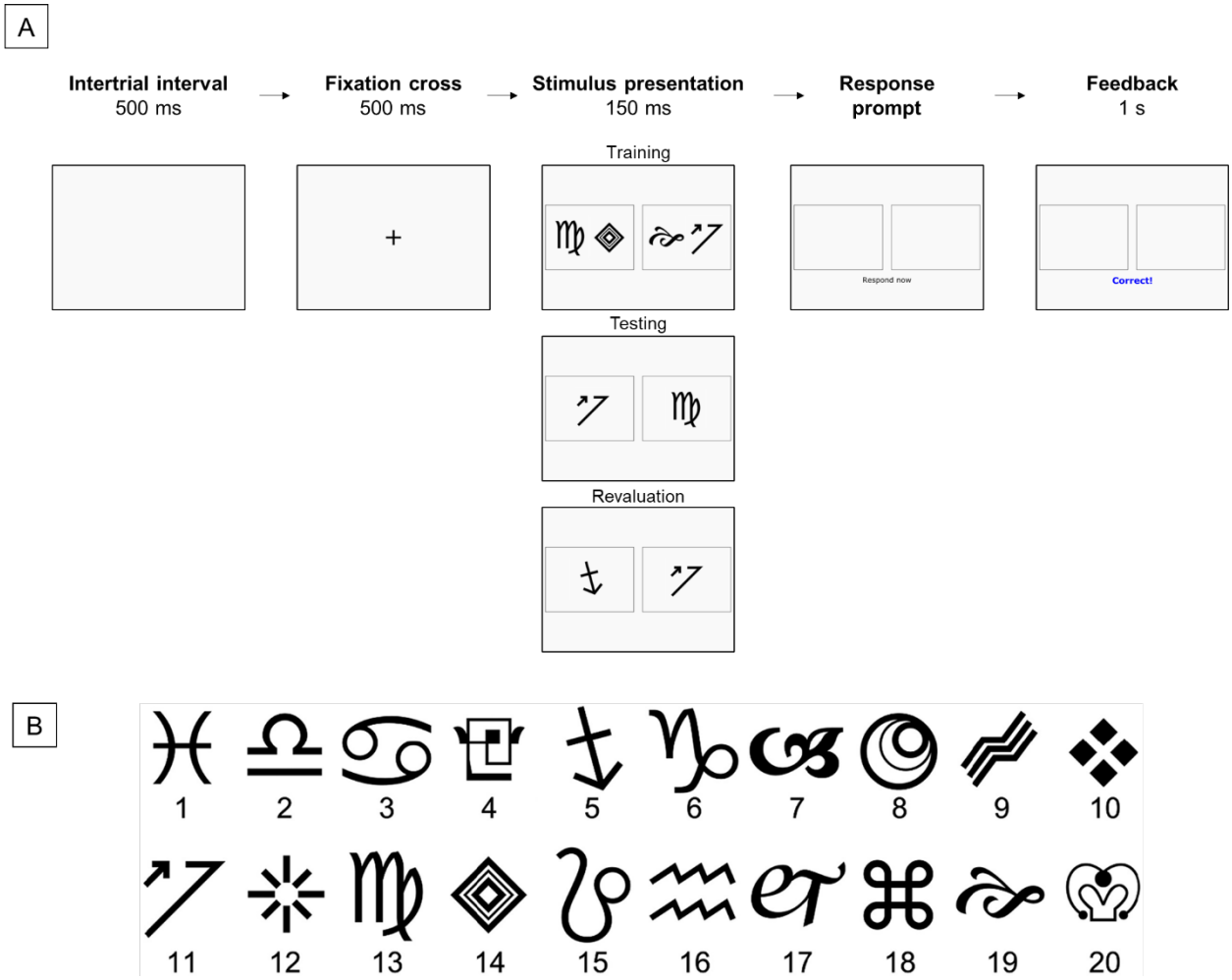


FIGURE 2 Mean percentage of choice before and after revaluation. The left panel shows mean percentage choice of the overselected (OS) and underselected elements of the revalued compound, and the right panel shows choice of the elements of the control compound. Filled circles show choice before revaluation, and unfilled squares show choice after revaluation. Error bars show the SEM.

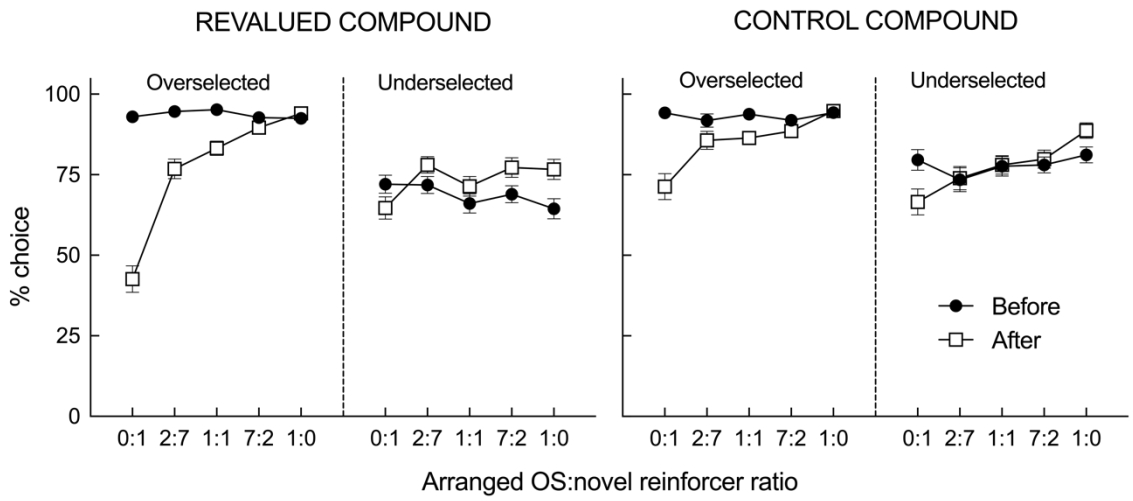


FIGURE 3 Obtained probability of reinforcers (A) and percentage of choice (B) during revaluation. OS = overselected element. Panel A: Shaded boxes show obtained reinforcer probabilities for the overselected element, and unshaded boxes show probabilities for the novel element. Each box represents the interquartile range, and the horizontal line inside each box shows the median. The whiskers show the minimum and maximum obtained reinforcer probabilities. The thick, solid horizontal bars show arranged probabilities. The arranged OS: novel reinforcer ratios are shown in parentheses for each group. Panel B: Percentage choice of the overselected (OS) element as a function of the probability of reinforcers for that element (*p*[rf for OS]). Error bars show the SEM.

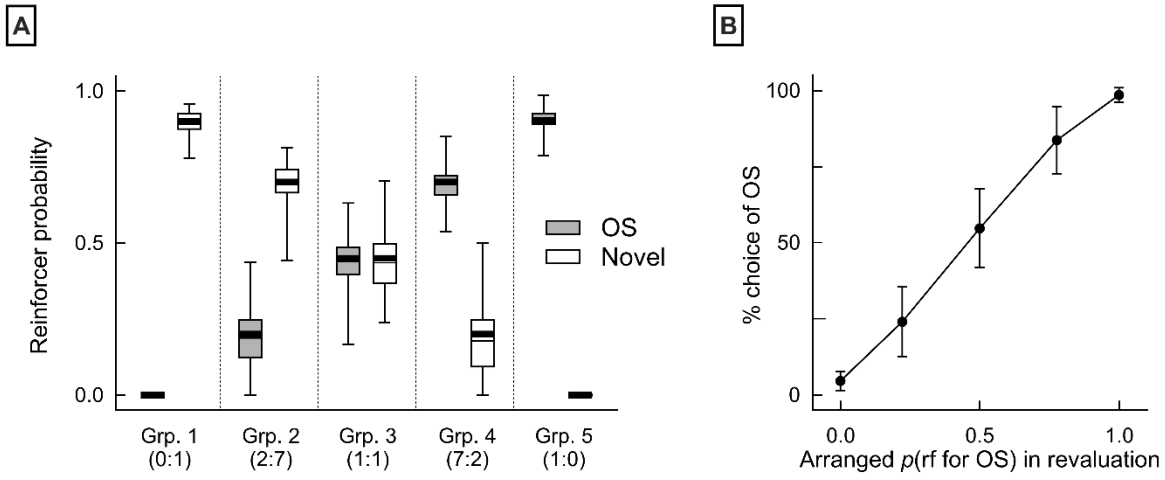
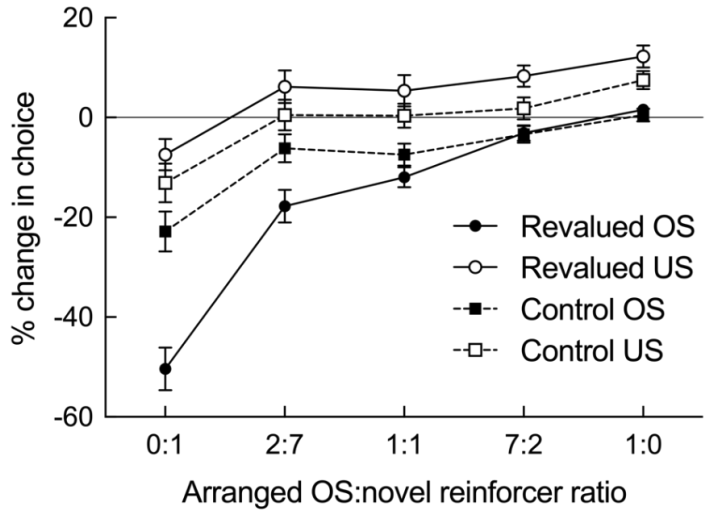


FIGURE 4 Mean percentage of change in choice. Circles show data for the revalued compound, and squares show data for the control compound. Filled symbols show changes in choice for the overselected (OS) element, and unfilled symbols show changes in choice for the underselected (US) element. Error bars show the SEM.



Appendix A: Individual participant data

FIGURE A1 Percentage of change in choice for individual participants in Group 1 for whom the arranged reinforcer ratio was 0:1 (overselected:novel) during revaluation. The solid horizontal line denotes the mean percentage of change in choice.

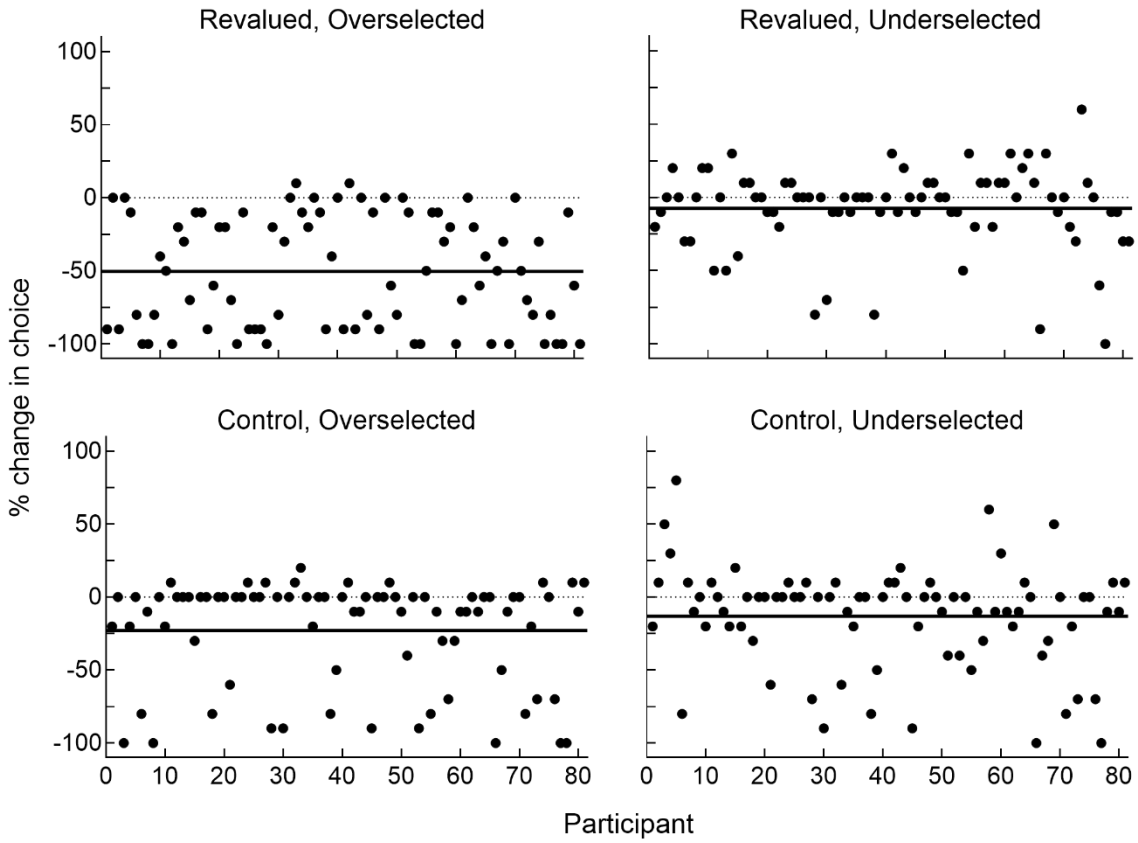


FIGURE A2 Percentage of change in choice for individual participants in Group 2 for whom the arranged reinforcer ratio was 2:7 (overselected:novel) during revaluation. The solid horizontal line denotes the mean percentage of change in choice.

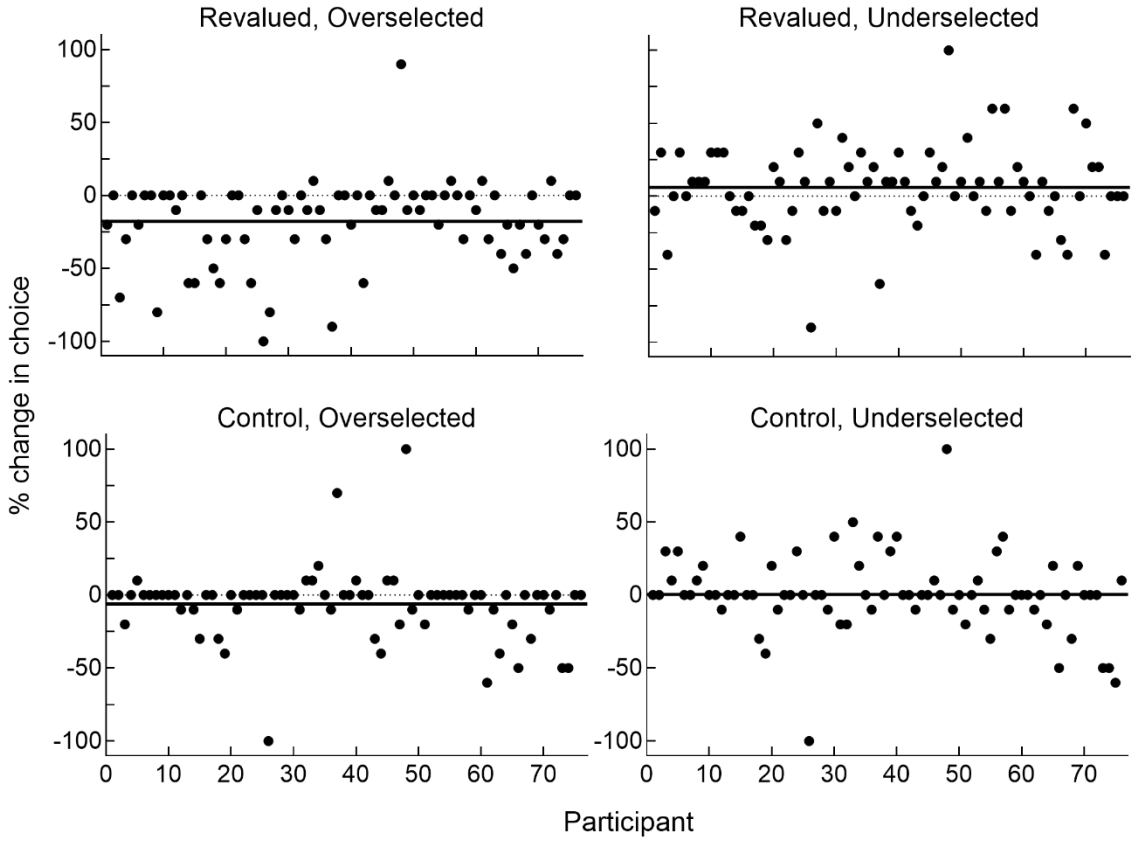


FIGURE A3 Percentage of change in choice for individual participants in Group 3 for whom the arranged reinforcer ratio was 1:1 (overselected:novel) during revaluation. The solid horizontal line denotes the mean percentage of change in choice.

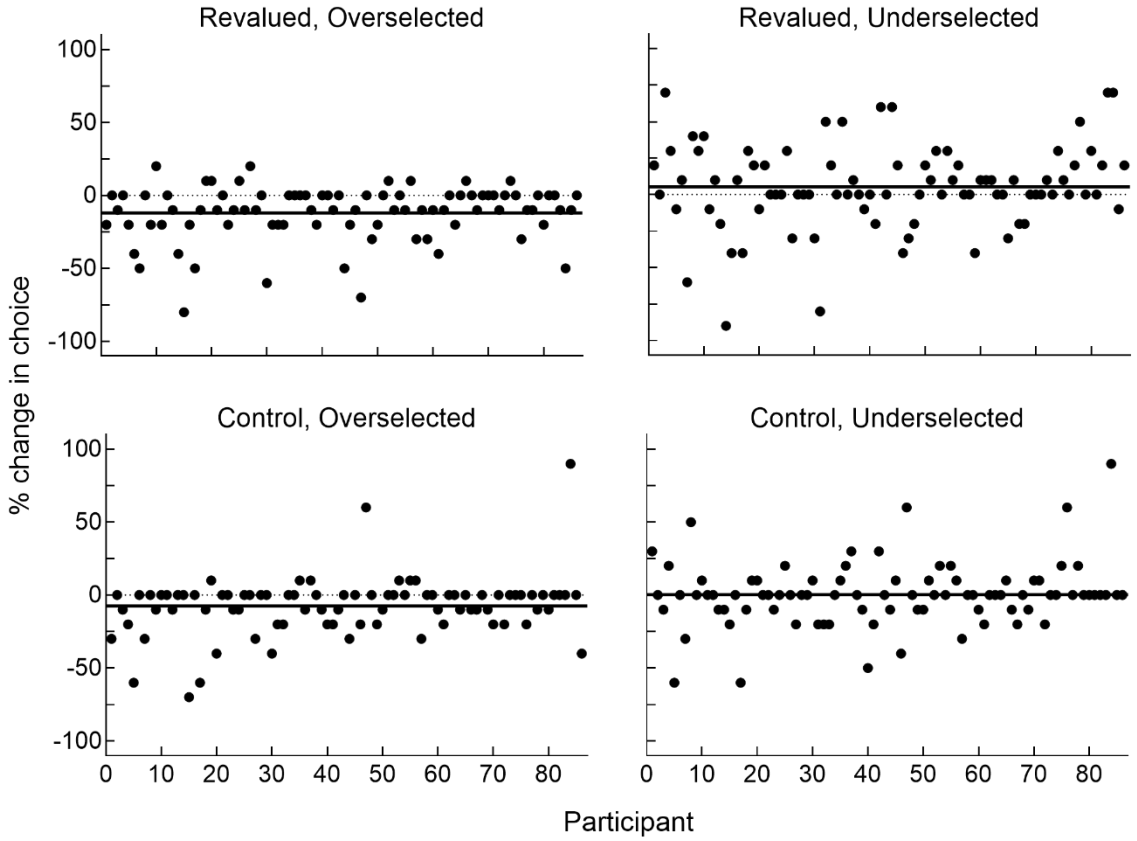


FIGURE A4 Percentage of change in choice for individual participants in Group 4 for whom the arranged reinforcer ratio was 7:2 (overselected:novel) during revaluation. The solid horizontal line denotes the mean percentage of change in choice.

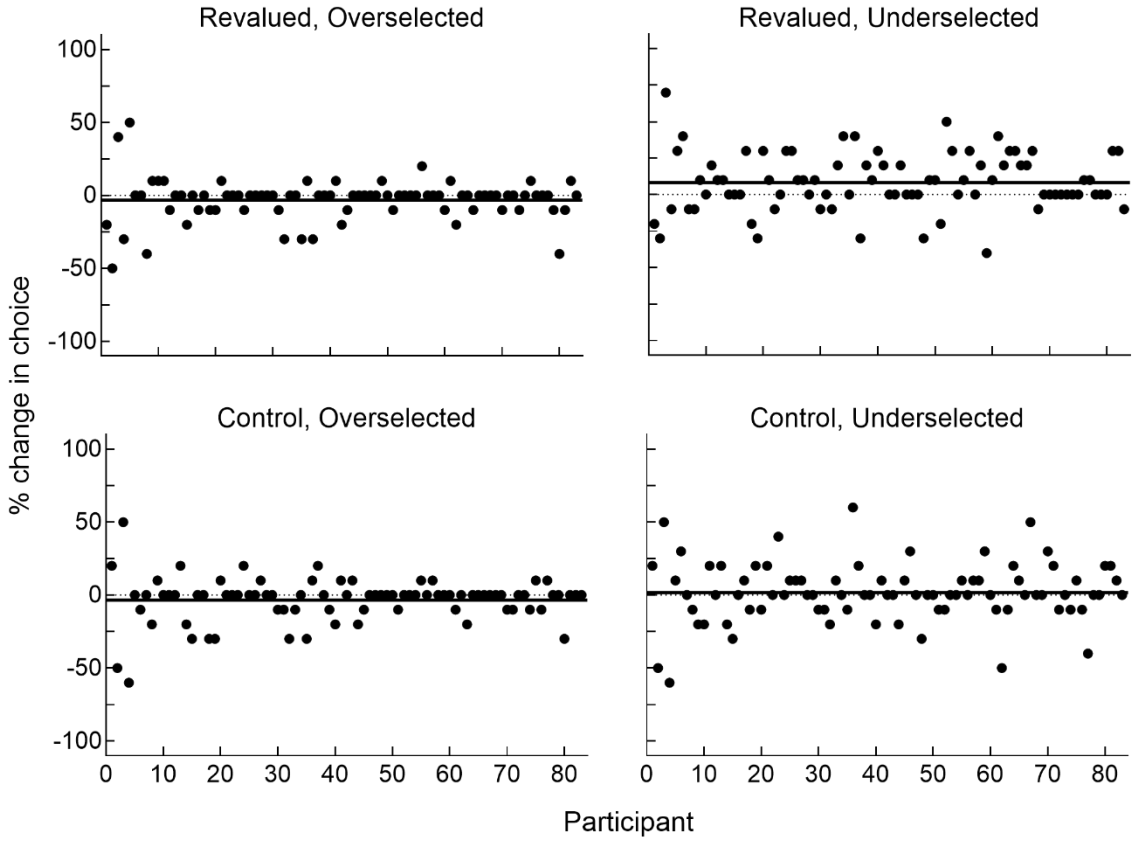


FIGURE A5 Percentage of change in choice for individual participants in Group 5 for whom the arranged reinforcer ratio was 1:0 (overselected:novel) during revaluation. The solid horizontal line denotes the mean percentage of change in choice.

