Behavioural Processes 158 (2019) 32–40

Contents lists available at ScienceDirect

Behavioural Processes

journal homepage: www.elsevier.com/locate/behavproc

Generalization of learned variability across multiple dimensions in humans

Xiuyan Kong[a,*], James S. McEwan[a], Lewis A. Bizo[a,b], Mary T. Foster[a,b]

[a] School of Psychology, University of Waikato, New Zealand
[b] School of Psychology and Behavioural Science, University of New England, Australia

ARTICLE INFO

Keywords:
Behavioral variability
Generalization
Human
U-value

ABSTRACT

This study examined whether trained variability would generalize across dimensions of the target response. Two experiments used a computerized rectangle drawing task that required participants to click and drag a mouse cursor to create rectangles on a computer screen. In Experiment 1, one group received points when successive rectangles varied in their size, shape and location (VAR), another group were yoked to the VAR group and received points that were allocated to them using a yoking procedure (YOKE), regardless of the variability in the size, shape or location of the rectangle drawn. Variability was higher for a dimension when variability on that dimension was directly reinforced. In Experiment 2, three groups of participants received points when rectangles varied on two dimensions; each group differed in the two dimensions that required variation. Variability was again higher for the reinforced dimensions for two of the three groups. Comparison with the YOKE group showed that the variability on those dimensions where variability was not directly reinforced was affected by reinforcement for variability on the other dimensions. Specifically, the variability in Shape and Location was significantly higher when these two dimensions occurred with other dimensions where variability was reinforced (as in Experiment 2) compared to when they were not required to vary (as in the YOKE group). This suggests that, for these two groups, the reinforced variability on the other two dimensions generalized to the third dimension. Implications of this finding to our understanding of factors that promote behavioral variability are discussed.

1. Introduction

The generalization of stimulus control over responses that occurs across stimuli and contexts has both theoretical and applied significance. In applied settings, generalization is usually considered to be a key criterion for judging the success of an intervention (Arnold-Saritepe et al., 2009; T. F. Stokes and Baer, 1977; T. F. Stokes and Osnes, 1989). It has previously been demonstrated that behavioral variability is sensitive to the contingencies of reinforcement and can come under discriminative stimulus control (Denney and Neuringer, 1998; Neuringer, 2002; Neuringer and Jensen, 2013; Page and Neuringer, 1985).

One question that is of both theoretical and applied interest is whether trained variability will generalize to new contexts. One of the suggestions for promoting generalization in applied settings is to train diverse responses to increase the likelihood of responses contacting reinforcement outside of the training environments (e.g., T. F. Stokes and Baer, 1977; T. F. Stokes and Osnes, 1989). Reinforcing varying aspects of the response during training can be seen as a step to the training of diverse responses. In applied settings, training of response variability does not usually precede training the prerequisite behaviors or responses– such training would fit with "training diverse exemplars" as described by T. F. Stokes and Baer (1977). Schmidt and Bjork (1992) suggested that training responses to vary could facilitate the generalization of such behavior to novel contexts.

There have been a limited number of studies that have investigated the extent to which learned variability would generalize across behaviors or contexts. This question is important because if a child was trained to vary in the number of colours they used when painting, might we also see some increase in variability in their use of building blocks or other forms of creative play? Some insight into an answer to that question is provided by P. D. Stokes et al. (2008) who trained participants using a computer based task to move a lit box from the top of a pyramid, shown on a screen, to different endpoints at the bottom of the pyramid. They reported that requiring variability in the paths taken to reach designated endpoints for reinforcement resulted in participants using more different paths than when this was not the case. When tested, after training, on a different pyramid and on endpoints that had

* Corresponding author.
E-mail addresses: kitt_kong@hotmail.com (X. Kong), james@behavioursolutions.nz (J.S. McEwan), lbizo@une.edu.au (L.A. Bizo), mary@behavioursoluitons.nz (M.T. Foster).

https://doi.org/10.1016/j.beproc.2018.10.020
Received 22 March 2018; Received in revised form 18 October 2018; Accepted 29 October 2018
Available online 01 November 2018
0376-6357/

never been trained on the original pyramid, those participants reinforced for varying paths remained more variable with the new pyramid and to the new endpoints than those that had not been reinforced for varying (P. D. Stokes et al., 2008). Thus, after experiencing the reinforcement contingency that required variability, the trained level of variability generalized to a different (but similar) task and to new elements of the same task.

Evidence of generalization of learned variability also comes from studies in the exploratory behavior of rats. Weiss and Neuringer (2012) trained a group of Long-Evans strain (LE), considered “bold” as they tend to spend relatively long times in the centre of a space and would rear on their hind legs more to inspect their surrounding environment, and a group of Piebald Virol Glaxo (PVG) rats, considered “shy” as they tend to spend little time in the centre of a space and rear little. These groups received food pellets for varying their responding. Two other groups of rats of the same strains - sisters of the experimental groups - received food pellets yoked to their sisters, independent of response variability. Both the LE and the PVG rats that were reinforced for varying their responses spent more time interacting with the objects and used more variable ways of interacting with the objects - ways that were less commonly used by other rats interacting with the objects, than their sisters in the yoked condition in which variability was not required for reinforcement. The level of variation in the responses from the genetically “shy” PVG rats in the variability group came to reach a level that was the same as the genetically “bold” LE rats. When tested for generalization of the trained level of variation in responding with a completely different environment and with a set of new novel objects that they had not encountered previously, in which pellets were hidden, the rats that had received reinforcers for variable interactions successfully found and consumed more food pellets in the new environment than were found by the other groups.

These findings demonstrated that the increased level of variability from training generalized to new environments with new objects. The absence of the training resulted in less exploration in the new environment and less different ways of interacting with new objects which in turn lowered the probability of earning reinforcers. This finding has significant implication for intervention for individuals with low response rates, such as for depressed individuals.

There is also evidence of possible generalization of behavioral variability from two studies that did not directly investigate this. Maes (2003) studied the effect of contingent and non-contingent reinforcement on response variability in three response sequences. Participants were asked to generate three-digit sequences that consisted of combinations of the numbers 1, 2 and 3, for example 111, 211, or 321. One of the experiments included two groups of participants; one group first received contingent reinforcement on variability of the sequences (VAR1) followed by reinforcement that was yoked to the number and distribution of reinforcers received previously in VAR1 but independent of variability of current response (YOKE1). Participants from the other group first received reinforcement that was yoked to the VAR condition of the first group, but independent of the variability of the current sequences (YOKE2), followed by reinforcement contingent on the variability of the sequences (VAR2). It was found that the percent of sequences that would have met the variability requirement was higher for YOKE1 condition when it followed the variability required condition (VAR1); however, the percent of sequences that would have met the variability requirement was lower for the YOKE2 condition when it preceded the condition where variability was required (VAR2). The results demonstrated that after participants learned to vary under contingent reinforcement for variability, their tendency to vary was greater when reinforcement was no longer contingent than it was for participants who had no previous experience of being reinforced to vary and only received non-contingent reinforcement.

Some evidence for generalization was also reported by Souza et al. (2010) using a task similar to Maes’s (2003) three response sequence generation task. In one phase of a condition in Souza et al. (2010) second experiment, participants received reinforcers probabilistically (50%) and independently of the variability of their sequences. Performances of two groups of participants were compared. One group experienced the variability independent phase after a phase in which variability of the sequences was required for reinforcement. The other group experienced the variability independent phase prior to the variability phase. Souza et al. found that participants produced more variable sequences in the variability independent phase if this phase had been preceded by a phase in which variability was reinforced. Therefore, similar to results found in Maes (2003), having experienced contingent reinforcement on varying the sequences positively influenced the variability of sequences even when it was not required for reinforcement. Results from both these studies suggest that the acquired level of variability from contingent reinforcement generalized to a condition where variability was not directly reinforced.

Generalization of acquired variability has also been investigated in applied settings with individuals with autism. Lee et al. (2002) trained three individuals with autism to use more variable and appropriate responses to questions from adults. The number of appropriate responses started low for all three before increasing greatly for two of them during the intervention when a Lag 1 schedule was used. The number of novel verbal statements observed for these two individuals also increased. Later, when reinforcement was no longer in effect, the same two individuals continued to use more variable and appropriate responses in the presence of new therapists in the same setting and in the presence of the same therapist in a different setting. Lee et al.’s study demonstrated that the level of variability in functional verbal responding in individuals with autism acquired under training could generalize across people and settings. Miller (2012) examined if variability in shapes built with wood blocks or patterns created using pegs on a foam board trained under a Lag 3 schedule would generalize to variability in the sequence of colors used in painting in two squares on a paper with children with autism. He concluded that the stimulus control of the variability, and the repeat contingencies used was established for shapes or patterns, did not generalize to the new task.

Generalization of learned variability has been investigated in creativity studies. In one study, two children received descriptive social reinforcement for painting different forms while they received no reinforcement for a block-building task in the same experimental session (Goetz et al., 1973). The number of different forms in the paintings increased when the reinforcement contingency was in place for both children; novel forms were more likely during these sessions. Variability of forms used in the block-building task was never reinforced; however, the number of different forms built in each session increased across the sessions when variability of shapes in the painting was reinforced. Training variability in responding in one task increased the variability in another topographically different task that occurred in the same session; thus generalization was observed.

Holman et al. (1977) reinforced form diversity in the shapes of felt-pen drawings by three children and then assessed the form diversity on three other tasks completed on the same day but with no reinforcement. These three tasks included one topographically similar, easel drawing and two topographically different tasks, wooden block and Lego™ building. For felt-pen drawing the number of different forms increased when the reinforcement contingency was in place for all three children; so did the number of new forms for two children. For easel painting, there was a clear increase in the number of different forms observed for one child, a reasonable increase for another, and little increase for the third. The number of new shapes accumulated over sessions also increased for all children. Thus, trained variability generalized from the felt-pen drawing to easel painting. For block-building and Lego™ building, there was no increase in the number of different forms built within a session nor in the number of new forms accumulated over sessions. Generalization of the trained variability seemed to have occurred for tasks that were topographically similar but not for topographically different ones.

The results of studies of the generalization of learned variability, both in experimental and applied settings, across tasks, and across contexts, are not consistent with creativity-related studies (e.g., Goetz et al., 1973; Holman et al., 1977), nor with studies with individuals with autism (Lee et al., 2002; Miller, 2012). However, there are some risks of confounds by other variables in the studies by Holman et al. and Miller. For example, Holman et al. (1977) commented that making an immediate judgement on whether the form built was different and whether it qualified for reinforcement was difficult and could result in errors. Miller used the instruction "make something' for all tasks; however, it is not clear if that instruction applies as well to painting as it might to other tasks. Other studies in controlled experimental environments (e.g., P. D. Stokes et al., 2008; Weiss and Neuringer, 2012), as well as studies that did not directly investigate generalization of variability (Maes, 2003; Souza et al., 2010), show more promising results.

Though generalization across behaviors and/or settings is the ultimate goal for training variability, it is important to test for generalization of trained variability in simplified experimental environments where confounding variables can be controlled. The present experiments used a computerised rectangle drawing task, similar to that used by Ross and Neuringer (2002), to examine whether learned variability would generalize across different dimensions of the response. Each response created a rectangle that had three measurable dimensions: Area, Shape and Location. In one experiment, one group of participants received reinforcement only when a response varied on all three dimensions; that is, they needed to create rectangles of different sizes, shapes and at different locations. Another group received the same amount of reinforcement for drawing rectangles without the requirement of varying any of the dimensions. This task allows for the measurement of multiple dimensions of a response that occur simultaneously, and so affords the ability to differentially reinforce variability in each dimension.

Ross and Neuringer (2002) found that the sizes, shapes and locations of the rectangles were more variable when reinforcement was contingent on variability in all dimensions of the response than when it was not. Ross and Neuringer's (2002) second experiment provided a reinforcer when participants varied their responses on two dimensions and repeated on the third. Variability was shown to be high for the dimensions that were required to vary and low for the dimension that was required to repeat. They demonstrated that reinforcement can influence multiple dimensions of a response simultaneously; and that it can also be used to produce different levels of variability over each of these dimensions. This rectangle drawing task provides an opportunity to examine whether reinforcing variability on some dimensions influences the variability of another dimension where variability is not directly reinforced.

Experiment 1 was a direct replication of Ross and Neuringer's (2002) first experiment in which participants received reinforcers when a rectangle drawn on a computer screen varied in three dimensions (Area, Shape and Location) from previous rectangles. Success in replicating their results would provide the basis for then studying the generalization of learned variability across dimensions in this same task.

2. Methods

2.1. Participants

Forty adults (23 females and 17 males), aged between 17 and 42 years, participated. They were recruited by advertisements posted on noticeboards around the University and on online course noticeboards. To facilitate the yoking design used in this experiment the first 20 participants recruited were assigned to the experimental group; and the next 20 were assigned to the control group. Participants who were first year psychology students were given 1% course credit. Ethics approval for this research was obtained from the University of Waikato School of Psychology Research and Ethics Committee (2010:36).

2.2. Apparatus

A PC computer with a 20-in monitor, a keyboard and a mouse were used to present instructions and record participants' responses. A computerised rectangle drawing task was created based on Ross and Neuringer's (2002) description of their task. Three dimensions – Area, Shape and Location – of the rectangles drawn were measured and monitored for variability. Area was defined by height times width; Shape was defined by the height to width ratio; and Location was defined by the (x, y) position of the centroids of the rectangles.

2.3. Creating 16 categories

A computer was programmed to randomly generate 500,000 rectangles within the confines of the computer screen; four measures were recorded for each rectangle: the sizes (Area), the ratio of width to height (Shape) and the x and the y coordinates of the centroid (Location). These four measures were then sorted independently of each other in an ascending order. Then, two of the lists, for Area and Shape, were divided into 16 sets with equal number of values (31,250 * 16 = 500,000); while the rest of the two lists, x and y values, were divided into four sets of equal number of values. For Area and Shape, the upper and lower bounds of each set for these two dimensions defined the 16 categories (described later) used for assessing variability. Using this method, combined with x and y data, 16 categories with similar sized regions on the screen were defined.

2.4. Procedure

Participants were tested individually. They were given time to read the information sheet and ask questions before the experimenter read the consent form to them. After signing the consent form, participants proceeded to the computer on which the instructions were displayed. The instructions were as follows:

> *"To play, simply click the mouse and drag on any diagonal to create a rectangle. Release the mouse button when you are satisfied with your rectangle. The object of this game is to get the most points. You have received points for your actions whenever you hear the ascending tones. There will be 300 trials. It will take approximately 15 min. Feel free to stop the experiment at any time. Please ask the researcher if you have any questions. Click anywhere on the screen to start."*

Light grey rectangles appeared on a dark grey background when the participants dragged the mouse diagonally to move the cursor on the screen. The experimenter left the room after the participant made two or three correct responses. Ascending tones (100-ms 1500-Hz followed by 100-ms 2000-Hz) occurred immediately after rectangles that met the reinforcement criteria were drawn. Every rectangle drawn remained on the screen for 500 ms and mouse clicks during this period had no effect; the screen cleared after 500 ms and the next trial started. Participants were debriefed at the end of the experiment.

2.5. Reinforcement of variability (VAR)

In order to encourage variability across responses, a threshold contingency was used. Reinforcement was given when the rectangle fell into categories that had been used less frequently in the recent past than other categories. Relative frequencies of occurrence were calculated for all categories for each dimension after each response by dividing the number of occurrences of a category by the total number of trials up to the current trial. For example, if Category 1 for the area dimension had been used 18 times over 300 trials, the relative frequency would be .06. For each rectangle drawn three categories, one from each dimension, would result. The relative frequencies of these categories from the immediately preceding trial were then compared to a threshold value to determine if the rectangle drawn would qualify for reinforcement.

For the experimental group, where variability on all three dimensions, Area, Shape and Location was required (VAR), participants were awarded a point only when the relative frequencies of the categories for all three dimensions were lower than a threshold value. An exponential weighting function was applied to trials that received points by multiplying all relative frequencies by .99 so that recent categories were weighted more than earlier ones. A detailed description and an example for the calculation of relative frequencies and the weighting procedure can be found in Denney and Neuringer (1998) and Ross and Neuringer (2002). These weighted relative frequencies were used to assess whether the categories resulting from a response were less frequent than the level specified by a threshold value. Ross and Neuringer (2002) suggested that responses of optimal variability would be similar to those seen in random responses. Thus, the threshold value was calculated based on what a random generator would do when drawing 300 rectangles. If the responses were evenly distributed across all 16 categories, the resulting threshold value would be .0625 ((300/16)/300). However, a threshold value of .0825 was chosen instead by Ross and Neuringer who found that this level of difficulty for this multi-dimensional task was appropriate after piloting different threshold values with humans. In the present experiment, we used the same threshold value as Ross and Neuringer (2002). Thus, for the VAR group, in order to gain a reinforcer (points), participants had to draw rectangles of which the weighted relative frequencies (WRF) of categories for Area, Shape and Location dimensions were lower than .0825. In other words, participants had to vary all three dimensions of the rectangles to get reinforcement.

2.6. Reinforcement independent of variability (YOKE)

A yoking procedure was used to guarantee that participants in the control group (YOKE) received the same number of reinforcers, at the same time as the experimental group. Each participant in the YOKE group was paired with a participant in the VAR group. The trials on which reinforcers were delivered to a participant in the YOKE group were identical to that of their assigned partner in the VAR group. For the participants from the YOKE group, reinforcers were delivered independent of the types of rectangles they drew.

3. Measures

3.1. Number of trials meeting variability criteria

For the VAR group, the number of trials that met the variability criteria equalled the number of reinforcers they obtained in the session. For the YOKE group, although the reinforcers were not given based on the variability of their responses, whether the responses would have met the variability criteria if they were in effect was recorded.

3.2. U-values

U-value, a measure of overall variability across all responses without regard to the reinforcement contingency, was calculated separately for the three dimensions. U-value was calculated according to the following formula:

$$U - value = -\sum_{i=1}^{\beta} \frac{\alpha_i \times \log(\alpha_i)}{\log(\beta)}$$

In the formula β equals the number of possible categories (16 in this case) and α equals the relative frequency of category *i*. The U-value shows how equally the responses were distributed over categories, and can range from 0 to 1, where a value of 0 indicates strict repetition and 1 indicates responses are equally distributed to all categories. By way of an example, if one examines the domain of Area, if a participant drew 300 rectangles of Category 4, the U-value for area dimension would be 0; if the participant drew 18 rectangles in each of the 16 categories, the U-value will be 1. Higher U-values reflect greater variability in the responses across categories (see Kong et al., 2017 for a discussion of U).

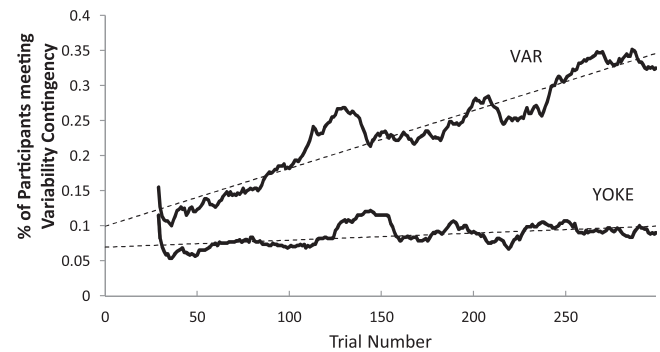

Fig. 1. The moving average across 30 trial blocks of the percentage of participants in the VAR and YOKE groups that met the variability criteria during Experiment 1. The solid lines are the moving average and the dotted lines are lines fitted by the method of least squares.

4. Results and discussion

4.1. Meeting reinforcement criteria

A one-way ANOVA across the two groups on the number of trials that met the reinforcement criteria showed the number was significantly higher for the VAR group (M = 70.6, SD = 25.85) than it was for the YOKE group (M = 27.25, SD = 26.25), $F(1, 39) = 27.69$, $p < .001$, partial $\eta^2 = .65$. This finding is consistent with Ross and Neuringer's (2002) results.

The percentage of participants in each group whose rectangles met the variability criteria on each of the 300 trials are shown in Fig. 1. The solid lines have been smoothed using a moving average across 30 trials. Dotted lines represent the linear regression for each group. On the first trial all participants' rectangles met the criteria to earn a point because the relative frequencies of all categories were zero at the start of the experiment. As they completed more trials, the percentage of participants in the VAR group whose rectangles met the criteria initially decreased markedly before increasing. Participants in the VAR group increasingly met the reinforcement criteria (varying their responses on all three dimensions) as the session progressed, while the percentage remained unchanged for the YOKE group which is also consistent with findings reported by Ross and Neuringer (2002).

4.2. U-values

U-values were calculated separately for each of the three dimensions based on the 300 responses for each participant. Mean U-values for each dimension for both groups are presented in Fig. 2. A mixed-design repeated-measures ANOVA was conducted, comparing the U-values over the groups (VAR vs YOKE), with dimension as the repeating factor. Mauchly's Test showed the data did not violate assumptions of sphericity, $W(2) = .960$, $p = .47$, and so no correction for this was required. The ANOVA showed that there was a significant main effect of group (VAR vs YOKE), $F(1, 38) = 27.88$, $p < .001$, partial $\eta^2 = .423$. There was a significant effect of dimension, $F(2,76) = 6.69$, $p = .002$, partial $\eta^2 = .150$, but the interaction between group and dimension was not significant, $F(2, 76) = 1.93$, $p = .153$, partial $\eta^2 = .048$. Post-hoc t-tests without corrections, comparing U-values for each dimension between the VAR and YOKE groups, revealed that mean U-values for the Area were significantly higher for the VAR group (.95, SD = .047) than

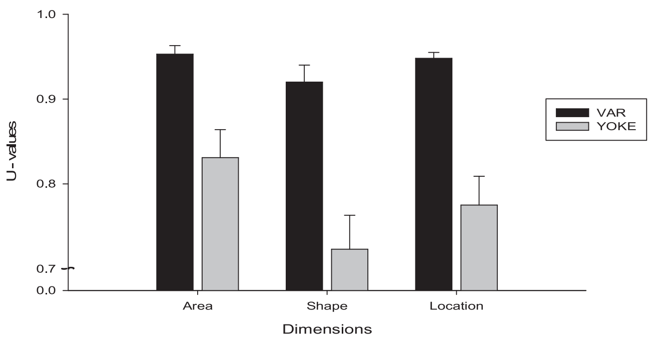

Fig. 2. Mean U-values for each dimension for the VAR and YOKE group. High U-values indicate high variability. Error bars show standard errors. Values below .7 have been omitted from the y-axis.

for the YOKE group (.83, $SD = .146$), $t(38) = 3.57$, $p < .001$; mean U-values for Shape were significantly higher for the VAR group (.92, $SD = .088$) than for the YOKE group (.72, $SD = .181$), $t(38) = 4.39$, $p = .012$; and mean U-values for Location were significantly higher for the VAR group (.95, $SD = .33$) than for the YOKE group (.78, $SD = .153$), $t(38) = 4.939$, $p < .001$

These results were consistent with results from Ross and Neuringer's (2002). The differences between VAR and YOKE groups' data for each of the three dimensions observed in the current study, however, are larger than those reported in Ross and Neuringer's (2002); and there were greater differences between mean U-values among the three dimensions for the YOKE group in the current study (between .12 and .21) than in Ross and Neuringer's (2002) study (within .01). In addition, however, a repeated measures ANOVA revealed that there was a significant difference in the U-values among the three dimension for the YOKE group, $F(2,38) = 4.55$, $p = .017$, partial $\eta^2 = .193$. Specifically, pairwise comparisons among the three dimensions revealed that U-values for Area were significantly higher than for Shape. This suggested that participants from the YOKE group had varied on the sizes (Area) more than on the shapes of the rectangles. No significant difference was found among U-values for the three dimensions for the VAR group.

The significant difference found between dimensions for the YOKE group suggests that the reinforcement delivered for this group, although independent of the variability of any dimensions, impacted the variability on the three dimensions differently. Specifically, there appeared to be a tendency for participants in the YOKE group to vary the sizes (Area) of the rectangles. When data for all participants were inspected, it was found that, for all three dimensions, most of the U-values for the VAR group were higher than those from the YOKE group. Where the scores were over 80, the differences in U-values between the VAR and YOKE pairs for the Area dimensions appear to differ from those for the same pairs for the Shape and Location dimensions. For example, participants from the YOKE group appear to have varied the area of the rectangles when they received more reinforcement (> 80) even though reinforcement was independent of the variability in the area of the rectangle they drew. This might have resulted from participants erroneously or superstitiously attributing reinforcement to their varying the area of the rectangles The same did not happen to variability of the Shape and Location dimensions, and this may suggest that when the reinforcement contingency was unclear to the participants, they tended to vary on the Area dimension rather than Shape and Location dimensions – showing a possible response bias toward varying on that dimension.

Experiment 1 successfully showed that higher variability in responses over multiple simultaneous dimensions can be obtained by reinforcement, thus replicating the results reported by Ross and Neuringer (2002). An important question remains, and that is whether learned variability generalizes across dimensions of the same behavior. To achieve this, one has to assess whether the variability in a dimension, where variability was not reinforced, changes as the result of reinforcing variability in other dimensions of the same behavior. In Experiment 2, a schedule that reinforced variability on only two dimensions was employed. Variability on the third dimension was allowed but it did not produce reinforcement. It should be noted, however, that the non-orthogonality of the three dimensions could be a confounding factor as variations in one dimension might be restricted by the parameters of the other dimensions. By way of explanation as to how this could impact our results consider the following examples – for rectangles drawn on at the edges or corners of the screen the variability of the size of those rectangles will be constrained compared to rectangles drawn in the centre of the screen, or if extremely large rectangles were drawn, the capacity of participants to vary in their selections of locations will be more constrained than if they drew small rectangles where the centroid of the rectangle could vary over a larger range. This non-orthogonality needs to be considered when examining results.

Results from Experiment 2 were combined with results from Experiment 1 to examine generalization across dimensions. In the previous experiment, for one group of participants (YOKE), variability in none of the three dimensions was required for reinforcement. In Experiment 2, reinforcement would be arranged to be contingent on variability of two of the three dimensions. Comparisons of the variability in a dimension that was not required to vary between Experiment 2 and the YOKE group from Experiment 1 facilitated the examination of generalization.

5. Method

5.1. Participants

Sixty participants who had not participated in Experiment 1 participated. They were recruited by advertisements posted on notice boards around the University and online. Participants who were first year psychology students were given 1% course credit. They were assigned to one of three groups in the following manner: The first participant was assigned to the Non_Area group, the second to the Non_Shape group, the third to the Non_Location group, and the fourth to the Non_Area group and so forth until each group had 20 participants. Informed consent was obtained before each experimental session.

5.2. Apparatus

The same rectangle drawing task and test environment were used, as in Experiment 1; the only difference was in the reinforcement contingency.

5.3. Procedures

5.3.1. Reinforcement of variability

The threshold value against which the WRFs for the dimensions to be varied would be compared was the same as in Experiment 1 (.0825). To earn reinforcement – delivery of a point – the WRF of the categories for two out of three dimensions had to be lower than the threshold value regardless of the WRF of the category used on the third dimension. One group received reinforcement when they varied the categories used for the dimensions Shape and Location (Non_Area), the second group received reinforcement when they varied the categories used for the dimensions Area and Location (Non_Shape), and the third when they varied the categories used for the dimensions Area and Shape (Non_Location).

5.3.2. Experimental procedures

The experimental procedure was the same as in Experiment 1 except that there was no YOKE group – the data from the YOKE group from Experiment 1 were used in the results section as an assessment of the

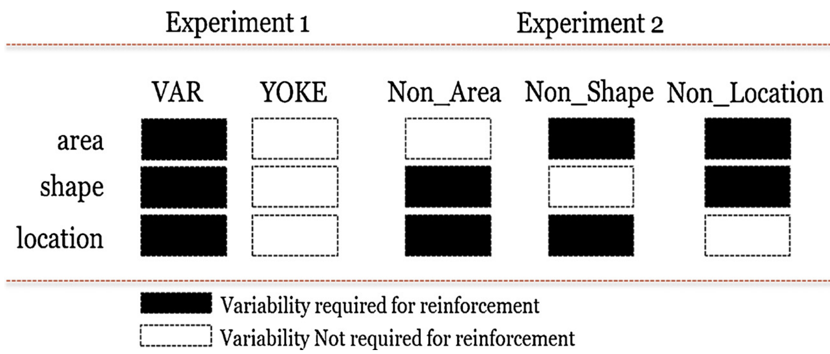

Fig. 3. Reinforcement arrangements. Filled boxes show dimensions that were required to vary for reinforcement; empty boxes show dimensions that were not required to vary for reinforcement.

levels of variability on the different dimensions when variability was not required for reinforcement. Planned comparison statistics were used to compare performance of the YOKE group from Experiment 1 with that of each of the groups from Experiment 2. Fig. 3 shows whether variability was reinforced for the three dimensions in each of the groups in Experiment 1 and Experiment 2. Dark boxes represent dimensions for which variation was required in separate conditions of Experiment 1 and Experiment 2. Comparisons were made of U-values for the dimensions where variability was not required (empty boxes); such as variability in the Area dimension between YOKE and Non_Area group, in the Shape dimension between YOKE and Non_Shape groups and in the Location dimension between YOKE and Non_Location groups. Comparisons of U-values across dimensions for each of the groups in Experiment 2 were also made to determine if the variability on the two dimensions required to vary was higher than the variability in the third dimension where there was no variability requirement.

The numbers of trials meeting reinforcement criteria were compared across the three groups to examine whether performances differed because of the different dimensions required to vary; any difference found would suggest a confounding effect of the dimensions on each other. The percent of trials with WRF less than .0825 was also calculated for each participant for each dimension as a measure of how variability changed over time.

6. Results and discussion

The U-values on each dimension from the YOKE group (Experiment 1) were compared to those from Experiment 2 to assess whether there was generalization of variability from the reinforced dimension to those dimensions where variability was not directly reinforced. Fig. 4 shows the mean U-values for each dimension from the YOKE group in Experiment 1 (empty bars) where all three dimensions were not required to vary for reinforcement. Also shown are the mean U-values for the Area dimension (slashed bar) from Non_Area group, for the Shape dimension (dark grey bar) from Non_Shape group and for the Location dimension (dotted bar) from Non_Location group. Overall, where variability was not required for reinforcement in Experiment 2, the mean U-values for all three dimensions are higher than for the YOKE group in Experiment 1.

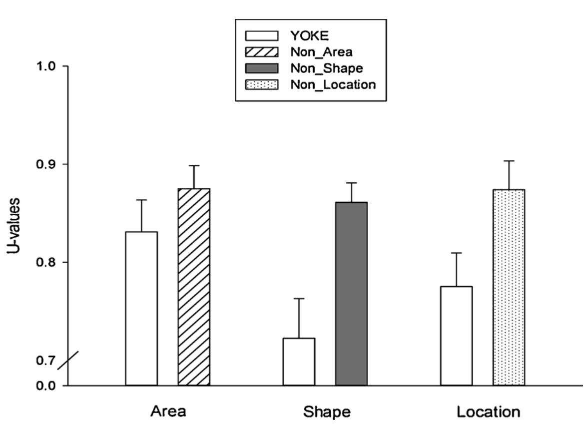

Fig. 4. Mean U-values for Area, Shape and Location dimensions when they were not required to vary across Experiment 1 and Experiment 2. Error bars represent standard errors. U-values below .7 were omitted from the y-axis.

Independent samples t-tests were carried out for the Area dimension between the YOKE and Non_Area groups, for the Shape dimension between the YOKE and Non_Shape groups and for the Location dimension between the YOKE and Non_Location groups. For the Area dimension, U-values were not significantly different for the Non_Area group and the YOKE group, $t(38) = 1.10$, $p = .281$, $r = .18$. For the Shape dimension, U-values were significantly higher for the Non_Shape group than for the YOKE group, $t(38) = 3.09$, $p = .004$, $r = .45$. For the Location dimension, U-values were significantly higher for the Non_Location group than for the YOKE group, $t(38) = 2.19$, $p = .035$, $r = .33$. Thus, the variability for the Shape and Location dimensions were confirmed to be significantly higher when these two dimensions occurred with other dimensions where variability was reinforced (as for the Non_Shape and Non_Location groups) compared to when they occurred with other dimensions that were not required to vary (as for the YOKE group).

These results suggested that for two of the three groups, Non_Shape and Non_Location groups, the reinforced variability on the two dimensions could have generalized to a third dimension where variability was not directly reinforced but occurred simultaneously with these two dimensions. These results also give support to the possible generalization of learned variability interpreted from studies by Maes (2003) and Souza et al. (2010). The lack of a difference in U-values produced by participants in the YOKE and Non_Area groups could be due to the relatively high U-values produced by the YOKE group; this was raised as a concern in Experiment 1 where U-values were significantly higher for the Area dimension than for the other dimensions for this group. Ross and Neuringer (2002) pointed out that the three dimensions are not completely independent because of the constraints of the computer screen but they suggested that the correlation among the three dimensions was extremely weak. Results from current experiments suggested, that even though weak, the correlation could have been strong enough to have impacted the data collected from our sample.

As task difficulty may have been different if the number of trials differed, the numbers of trials meeting the criterion for reinforcement were compared across the groups to determine if varying on any two of the dimensions for reinforcement affected the number of trials meeting the reinforcement criteria. A one-way ANOVA revealed that the numbers of trials (Non_Area $M = 114.10$, $SD = 39.96$; Non_Shape $M = 105.50$, $SD = 29.4$; Non_Location $M = 93.65$, $SD = 39.53$) did not differ significantly across groups, $F(2, 57) = 1.57$, $p = .217$. Therefore, while the groups had reinforcement contingent for varying with

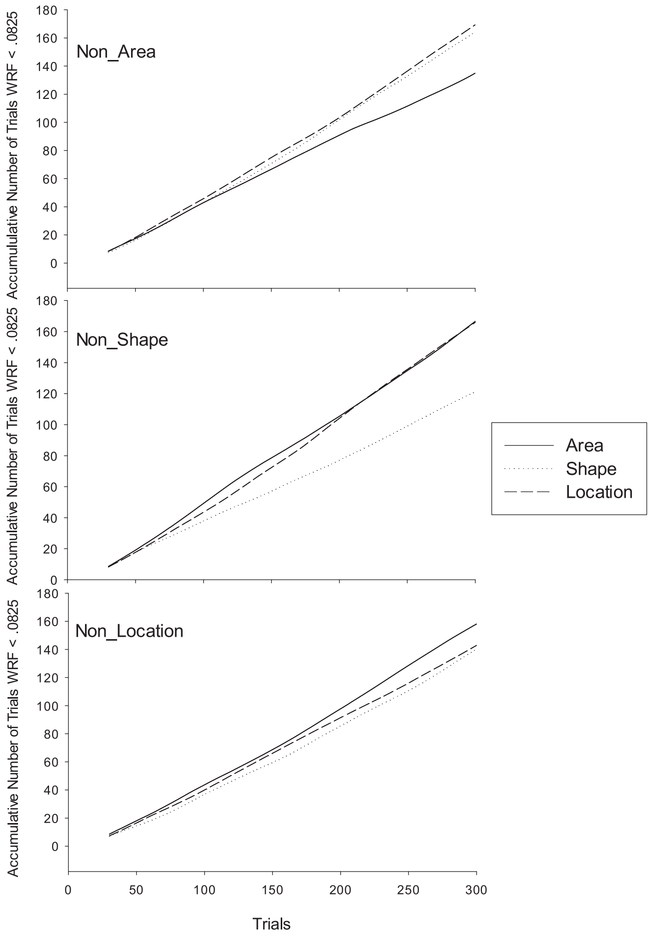

Fig. 5. Moving average of the cumulative number of trials WRF less than .0825 for Area, Shape and Location dimensions for Non_Area, Non_Shape and Non_Location groups.

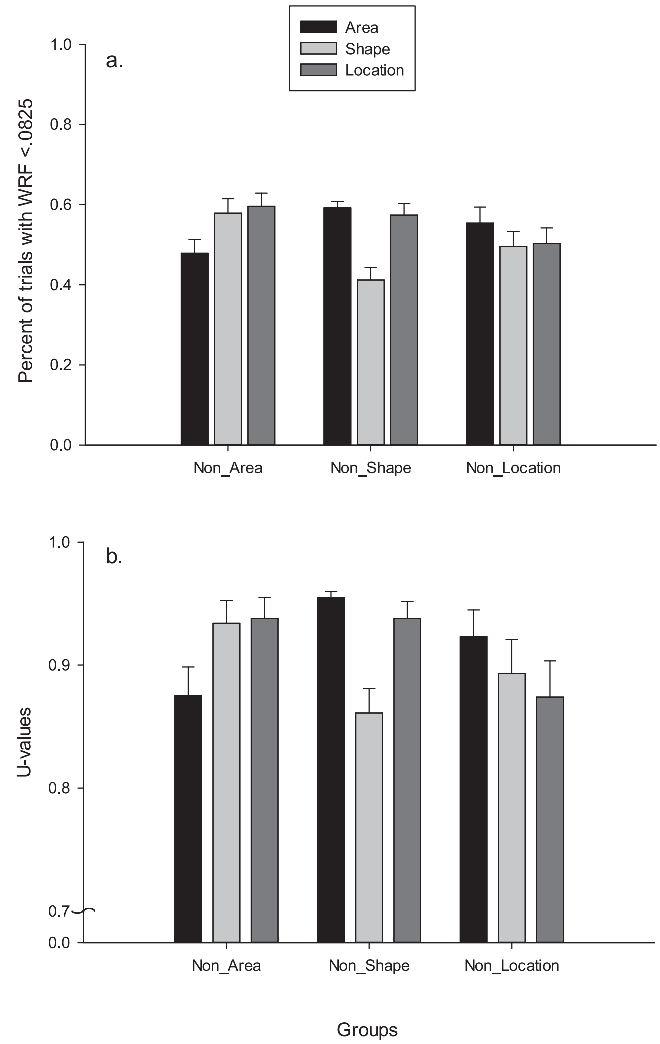

Fig. 6. (a) Percent of trials with WRF $<$ .0825 for Area, Shape and Location dimensions for Non_Area, Non_Shape and Non_Location groups. Error bars represent standard errors. (b) Mean U-values for Area, Shape and Location dimensions for Non_Area, Non_Shape and Non_Location groups. Error bars represent standard errors.

different dimensions, they met the criterion for reinforcement on a similar number of trials.

The numbers of trials with WRF $<$ .0825 were counted for each dimension separately for each of the three groups. This measure shows whether the variability of categories use differed across dimensions within each group. Fig. 5 shows the mean cumulative number of trials with WRF $<$ .0825 over the 300 trials for each dimension for the three groups. Data were smoothed by plotting the moving average of 30 trials.

As can be seen in Fig. 5, for the Non_Area and Non_Shape groups, there was a steeper increase for the two dimensions that were required to vary (Shape and Location for Non_Area and Area and Location for Non_Shape) than for the dimension that was not required to vary (Area for Non_Area and Shape for Non_Shape). However, the increase appeared to be similar for all three dimensions for the Non_Location group.

For ease of comparison, the percent of trials with WRF $<$ .0825 for each dimension for all three groups are plotted in panel a. Fig. 6. As can be seen, percent of trials with WRF less than .0825 is lower for the Area dimension than for the Shape and Location dimensions for the Non_Area group; and it is lower for the Shape dimension than for the Area and Location dimensions for the Non_Shape group. However, for the Non_Location group, the percent of trials with WRF less than .0825 for the Location dimension appears to be lower than that for the Area dimension only; the percentages appears to be equal for Shape and Location dimensions. This means that for the Non-Location group, although reinforcement was contingent on the variability of the Area and Shape dimensions, the variability in the Shape dimension does not appear to be higher than that of the Location dimension.

Repeated measures ANOVAs were used to compare change in the percent of trials with WRF $<$ .0825 across dimensions. For the Non_Area group, the percent of trials with WRF $<$ .0825 was significantly different across the three dimensions, $F(2, 38) = 5.69$, $p = .007$, partial $\eta^2 = .23$. Pairwise comparisons showed that it was significantly lower for the Area dimension ($M = 143.55$, $SD = 45.48$) than for the Shape dimension ($M = 173.65$, $SD = 48.64$; $p = .012$) and for the Location dimension ($M = 178.85$, $SD = 44.81$; $p = .009$); the percent of trials with WRF $<$ .0825 for the Shape and Location dimensions did not differ significantly ($p = .639$). For the Non_Shape group, Mauchly's test indicated that the assumption of sphericity had been violated, $x^2(2) = 9.06$, $p = .011$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. Percent of trials with WRF $<$ .0825 differed significantly across dimensions, $F(1.43, 27.23) = 23.95$, $p < .001$, partial $\eta^2 = .56$. Pairwise comparisons indicated that percent of trials with WRF below the threshold value were lower for the Shape dimension ($M = 123.60$, $SD = 40.97$) than it was for Area ($M = 177.75$, $SD = 21.36$; $p < .001$) or Location ($M = 172.35$, $SD = 38.69$; $p < .001$); however, it was not significantly different ($p = .584$) for Area and Location. For the Non_Location group, percent of trials with WRF $<$ .0825 was not significantly different across Area (M = 166.30, SD = 53.04), Shape (M = 148.7, SD = 49.06) and Location (M = 151.00, SD = 51.93), $F(2, 38) = 2.23$, $p = .122$.

Overall, when reinforcement was contingent on varying the Shape and Location dimensions (Non_Area), the rectangles participants drew varied more on the dimensions of Shape and Location than they did on Area. When reinforcement was contingent on varying the Area and Location dimensions the rectangles participants drew varied more on the dimensions of Area and Location than they did on Shape. However, when reinforcement was contingent on varying the Area and Shape dimensions, participants drew rectangles of similarly infrequently used areas, shapes and locations.

U-values were calculated independently for each dimension for each of the three groups; and they were compared between the dimensions that were required to vary and the dimension that was not required to vary within each group and across groups. Panel b. of Fig. 6 shows mean U-values for the Area, Shape and Location dimensions for the Non_Area, Non_Shape and Non_Location groups. The differences in U-value across dimensions were similar to the differences in number of trials with WRF $< .0825$; U-values tended to be lower for the dimension that was not required to vary than for the two which were required to vary for the Non_Area and Non_Shape groups.

One-way repeated measures ANOVAs were carried out for each of the groups. For the Non_Area group, Mauchly's test indicated that the assumption of sphericity had been violated, $x^2(2) = 17.58$, $p < .001$; therefore the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. U-values were found to be significantly different across dimensions, $F(1.23, 23.41) = 4.87$, $p = .031$, partial $\eta^2 = .20$. Pairwise comparisons showed that U-value for the dimension that was not required to vary, Area (M = .88, SD = .11), was lower than that for both Shape (M = .93, SD = .08; $p = .04$) and Location (M = .94, SD = .08; $p = .029$), which were required to vary. For the dimensions that were required to vary, Shape and Location, no significant difference was found ($p = .702$). For the Non_Shape group, Mauchly's test indicated that the assumption of sphericity had been violated, $x^2(2) = 13.64$, $p = .001$, therefore the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. U-values were found to be significantly different across dimensions, $F(1.31, 24.81) = 20.66$, $p < .001$, partial $\eta^2 = .52$. Pairwise comparisons confirmed that U-values for the dimension that was not required to vary, Shape (M = .86, SD = .09), was lower for both the Area (M = .95, SD = .02; $p < .001$) and the Location (M = .94, SD = .07; $p < .001$) dimensions, which were required to vary. For the dimensions that were required to vary, Shape and Location, no significant difference was found ($p = .257$).

For the Non_Location group, the difference in U-values across dimensions was not significant, $F(2, 38) = 2.88$, $p = .069$.

As can be seen in panel b. of Fig. 6, the U-values for the Location dimension for the Non_Location group appeared to be similar to the U-values for Shape dimension for the Non_Shape group and for the Area dimension for the Non_Area group; all of these dimensions were not required to vary. A one-way ANOVA confirmed that the differences in U-values across these dimensions was not significant, $F(2, 57) = 0.095$, $p = .9$. This suggests to us that that the level of variability for Location was as high as the other two dimensions; the lack of difference was due to variability in the dimensions that were required to vary not being high enough. This is consistent with our previous suggestion that non-orthogonality of the dimensions is potentially problematic when using this task to investigate the effects of reinforcement on behavioural variability.

For all the dimensions that were required to vary, U-values for the Area and Shape dimensions for the Non_Location group appeared to be lowest of the six. A one-way ANOVA across these six U-values showed no main effects, $F(5, 48.39) = 1.629$, $p = .170$; Leven's test showed that the assumption of homogeneity was violated (Leven statistic (5, 114) = 3.771, $p = .003$) therefore, the Welch's test statistics is reported instead. Post Hoc analysis using LSD indicated that mean U-value for the Shape dimension for the Non_Location group was significantly lower than for the Area dimension for the Non_Shape group, $p = .022$; no other difference in U-values across dimensions were significant.

Overall, U-values for the dimensions that were not required to vary were not significantly different across the three groups. In all three groups the non-reinforced dimension had the lowest U-value. For the dimensions that were required to vary, although the differences in U-values between these dimensions were generally not significant, U-values for the Area and Shape dimensions for the Non_Location group were slightly lower than those for the Shape and Location dimensions for the Non_Area group and the Area and Location dimensions for the Non_Shape group. Had the variability in the Area and Shape dimensions for the Non_Location group been slightly higher, the results for the Non_Location group would have been consistent with those of the other two groups. The lack of difference for this group, appears to be the result of lower variability for the dimensions that were required to vary (as compared to variability in the dimensions that were required to vary for other two groups), rather than of higher variability for the non-reinforced dimensions. Thus, when Location was not required to vary, participants varied less on the two dimensions that were required to vary. The requirement to vary the Location dimension was common to the other two groups, Non_Area and Non_Location, and may have made the reinforcement contingency more salient to the participants. This suggests that it could be the non-orthogonality of the three dimensions that affected the variability of individual dimensions.

7. General discussion

Generalization of behavioural variability was examined by comparing the levels of variability on different dimensions of a multi-dimensional response. In separate experiments reinforcement was contingent on variation of 3 dimensions of a rectangle drawing task or on two but not the third dimension of that task. When two of the three dimensions under examination, Shape and Location, were not required to vary then their variability was found to be higher when variability in both the two other dimensions was being directly reinforced (Experiment 2) than when it was not (Experiment 1). This suggests that the enhanced level of variability generalised from the trained dimensions to the untrained one. Although for one dimension, Area, such enhancement was not found to be significant. We would argue that this was likely due to non-orthogonality of the dimensions – put plainly – it appeared to be easier to vary the size than it was to vary the shape and location of the rectangle. Our results are consistent with P. D. Stokes et al. (2008) who found generalization of trained variability across the same dimension in different tasks; that generalization occurred also across different and relatively independent dimensions as we found in our investigation. Moreover, our results also give support to the idea that the increase in variability in sequences produced when variability was not directly reinforced from Maes (2003) and Souza et al. (2010) was the result of generalization of reinforced variability across dimensions of the response.

Interdependency across dimensions of a response is a major difficulty faced in the current project when examining the generalization of learned variability. This is likely to be faced also by other researchers intending to investigate the generalization of reinforced variability across dimensions of the same response because the dimensions will often correlate naturally. For example, reinforcers could be provided for running faster. Running faster will inevitably result in longer strides and higher foot lifts. These are both part of the running process and could change, not because of any reinforcement contingency on them specifically, but because of the contingency on the speed.

One line of research examines the distances between dimensions in psychological space so that one can examine how close or distant the dimensions of interest are (e.g., Soto and Wasserman, 2010).

Developing a better understanding the level of separability of response dimensions in a multidimensional response space is important for a full understanding of conditions and factors that promote generalization of behavioural variability. The distances between dimensions in psychological space would differ greatly from individual to individual based on their personal history; it is therefore also important to develop an understanding of reinforcement history

The present results add to the extensive experimental research into the effect of reinforcement on behavioral variability. These studies show consistently that variability can be controlled by reinforcement. However, as Neuringer (2004) pointed out that, although behavioral variability is important for some areas of everyday functioning, there has been little exploration of the application of these findings. This present study provides evidence that reinforced behavioral variability can generalize across dimensions and across tasks, and this should encourage more research into the applications of training variability and of the generalizability of such interventions. There has been growing research in increasing behavioral variability in autistic individuals by directly reinforcing variability in behavior, such as functional responding, to improve on one of the core deficits, stereotypy, in these individuals (Susa and Schlinger, 2012). Research into how interventions designed to increase variability in autistic individuals generalize across people and contexts are also important ; and generalization should be included as part of assessment for effective treatment.

There are several more areas in which the application of reinforcing variability would be highly valuable. One is in increasing the variability of the ways individuals and animals approach and solve problems. One study investigated this, showing that after training rats learned to explore novel objects to obtain more food; and also that these results of learning generalized to a new environment (Weiss and Neuringer, 2012). Hopkinson and Neuringer (2003) found that moderately depressed college students could learn to increase the variability in sequence generation; the level of variability obtained sustained later when reinforcement was withheld. While such sequence generation appears to be distant from everyday functioning, a computerized task that incorporates elements including target behaviors and the to-be-generalized behaviors that resemble daily life behaviors would be a step forward in taking behavior variability studies from experimental towards applied settings. Such a computerised task affords the opportunity to monitor variability in multiple dimensions of the target behaviors and to deliver reinforcement accurately and immediately for variation on some or all of those dimensions.

A further area in which research into reinforcing variability and its generalizability can be applied, is in training creative responses. Studies examining increasing creative responses in humans have reinforced the use of different colours and the production of different shapes (e.g., Fallon and Goetz, 1975; Goetz and Baer, 1973; Holman et al., 1977). As discussed in the introduction, the problems faced in these studies, such as the subjectivity in assessing the behaviours, and the difficulty of ensuring accuracy and immediacy of reinforcement delivery, are difficult to overcome. The current research showed that a computerized task can be used to study, not only the effect of reinforcement on variability across multiple dimensions, but also whether the reinforced variability generalizes to other dimensions. It is possible that the future studies into increasing creative responses could make use of computerised tasks that would allow for reinforcing and monitoring the variability in independent dimensions such as colours, shapes, textures and so on.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Declaration of interest: none.

Acknowledgements

We thank Andrew Malcolm for his great patience in the process of creating and refining the computer program used in this study. Andrew was also responsible for maintaining the data on the server. We also thank Alan Eddy who had provided immediate technical support so that the experiments ran smoothly. And last but not least, we thank all participants who had contributed their time to this project.

References

Arnold-Saritepe, A.M., Phillips, K.J., Mudford, O.C., De Rozario, K.A., Taylor, S.A., 2009. Generalization and maintenance. In: Matson, J.L. (Ed.), Applied Behavior Analysis for Children with Autism Spectrum Disorders. Springer, New York, pp. 231–258.

Denney, J., Neuringer, A., 1998. Behavioral variability is controlled by discriminative stimuli. Anim. Learn. Behav. 26, 154–162.

Fallon, M.P., Goetz, E.M., 1975. The creative teacher: effects of descriptive social reinforcement upon the drawing behavior of three preschool children. School Appl. Learn. Theory 7, 27–42.

Goetz, E.M., Baer, D.M., 1973. Social control of form diversity and the emergence of new forms in children's blockingbuilding. J. Appl. Behav. Anal. 6, 209–217.

Goetz, E.M., Jones, K., Weamer, K., 1973. The generalization of creativity "training" in easel painting to blockbuilding. Paper presented at the Annual Convention of the American Psychological Association.

Holman, J., Goetz, E.M., Baer, D.M., 1977. The training of creativity as an operant and an examination of its generalization characteristics. In: Etzel, B.C., LeBlanc, J., Baer, D.M. (Eds.), New developments in behavioral research: theory, method and application. Lawrence Erlbaum Associates, Hillsdale, N.J, pp. 441–471.

Hopkinson, J., Neuringer, A., 2003. Modifying behavioral variability in moderately depressed students. Behav. Modif. 27, 251–264.

Kong, X., McEwan, J.S., Bizo, L.A., Foster, T.M., 2017. An analysis of U-value as a measure of variability. Psychol. Rec. 67, 581–586.

Lee, R., McComas, J.J., Jawor, J., 2002. The effects of differential and lag reinforcement schedules on varied verbal responding by individuals with autism. J. Appl. Behav. Anal. 35, 391–402.

Maes, J.H.R., 2003. Response stability and variability induced in humans by different feedback contingencies. Learn. Behav. 31, 332–348.

Miller, N.D., 2012. Stimulus control and generalization of operant variability in the block play of children with Autism. The Ohio State University Unpublished thesis.

Neuringer, A., 2004. Reinforced variability in animals and people: implications for adaptive action. Am. Psychol. 59, 891–906.

Neuringer, A., 2002. Operant variability: evidence, functions, and theory. Psychon. Bull. Rev. 9, 672–705.

Neuringer, A., Jensen, G., 2013. Operant variability. In: In: Madden, G.J. (Ed.), APA Handbook of Behavior Analysis: Methods and Principles, vol. 1. American Psychological Association, Washington D.C, pp. 513–546.

Page, S., Neuringer, A., 1985. Variability is an operant. J. Exp. Psychol. Anim. Behav. Process. 11, 429–452.

Ross, C., Neuringer, A., 2002. Reinforcement of variations and repetitions along three indepedent response dimensions. Behav. Processes 57, 199–209.

Schmidt, R.A., Bjork, R.A., 1992. New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. Psychol. Sci. 3, 207–217.

Soto, F.A., Wasserman, E.A., 2010. Integrality/Separability of stimulus dimensions and multidimensional generalization in pigeons. J. Exp. Psychol. Anim. Behav. Process 36, 194–205.

Souza, A., Abreu-Rodrigues, J., Baumann, A.A., 2010. History effects on induced and operant variability. Learn. Behav. 38, 426–437.

Stokes, P.D., Lai, B., Holtz, T.M., Rigsbee, E., Cherrick, D., 2008. Effects of practice on variability, effects of variability on transfer. J. Exp. Psychol. Hum. Percept. Perform. 34, 640–659.

Stokes, T.F., Baer, D.M., 1977. An implicit technology generalization. J. Appl. Behav. Anal. 10, 349–367.

Stokes, T.F., Osnes, P.G., 1989. An operant pursuit of generalization. Behav. Ther. 20, 337–355.

Susa, C., Schlinger, H.D., 2012. Using a lag schedule to increase variability of verbal responding in an individual with autism. Anal. Verbal Behav. 28, 125–130.

Weiss, A., Neuringer, A., 2012. Reinforced variability enhances object exploration in shy and bold rats. Physiol. Behav. 107, 451–457.