Influences of graphical image pairing on product preference

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

The Master of Applied Psychology in Behaviour Analysis

at

The University of Waikato

by

XIAOSHA MENG

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2018

Abstract

Many advertisements appear to rely on higher-order respondent conditioning to alter preference for a product brand. Given that higher-order conditioning is difficult to achieve, there may be other processes, such as those involved in stimulus equivalence, at work. The current experiment paired six “brand” stimuli with three different types of pictures. Three brands were paired directly with the pictures and the other three paired indirectly, using a stimulus equivalence procedure. The three types of pictures were “negative” (typical consequences of ingesting excessive sugar, such as heart disease, tooth decay, and diabetes), “neutral” (pictures of geometrical shapes) and “positive” pictures (pictures related to health and fitness). This research used a respondent procedure, which is more similar to real marketing methods than operant procedures, as used in previous research in this area. A preference assessment was conducted to evaluate transfer of function from the picture to the brands by investigating whether the pairings influenced participants’ preference for branded soft drinks. The results showed that negative pictures shifted preference more than positive pictures and direct pairing was more effective than an indirect pairing procedure.

Table of Contents

List of Figures

List of Tables

Advertising appears to rely heavily on classical conditioning, but classical conditioning with humans seems to be enriched/complicated by verbal abilities. Understanding the relationship between classical conditioning and verbal processes, such as those that appear to be occurring in stimulus equivalence procedures, may shed light on the mechanisms in response to successful advertising and counter-advertising.

Classical conditioning and advertising

Classical conditioning, or respondent conditioning, is a learning mechanism which was firstly studied systematically by Ivan Pavlov. He found that, after pairing a neutral stimulus (such as a tone (CS)) with food (US) for several times, presenting the CS alone will induce the salivation of the dog. Pavlov called the salivation in response to the CS as a Conditioned Response (CR). Classical conditioning had long been considered to be important in advertising, but empirical support is limited (Pornpitakpan, 2012). Most advertisement depends on patterns of second-order conditioning (or higher-order conditioning). In second-order conditioning, a new conditioned stimulus (CS2) is paired with a previously conditioned stimulus (CS1). In this process, CS1 serves as the US for conditioning CS2. For example, pair a bell (CS2) with the sound of the previously conditioned tone (CS1) should eventually result in salivating of the dog (CR) when it hears the bell. However, Pavlov found that, as the conditioning response extends beyond the second order, the strength of the CR becomes very weak. Since that time, many researchers have tried to extend respondent conditioning effects to higher orders (beyond second-order) but most have failed (Tryron & Cicero, 1989; Cicero and Tryon, 1989).

Staats and Staats (1957) conducted a study to examine whether a nonsense syllable could have the same function as a meaningful (i.e., already conditioned) word after the two had been paired multiple times. They paired words, which they described as positive, such as "pretty", "healthy" and "sweet," or negative, such as "worthless", "ugly" or "dirty," with nonsense syllables. The nonsense syllables were presented visually on the screen with a projector and, simultaneously, the examiner read the words loudly, and the participants were asked to repeat the words immediately after the examiner finished pronouncing it. Each nonsense syllable was paired four times with its corresponding word. After some initial orientation training, six nonsense syllables, YOF, LAJ, XEH, WUH, GIW, and QUG, were presented. Participants were divided into two groups. For the first group, YOF was paired with

positive words and XEH was paired with negative words, for another group, the function of YOF and XEH reversed. The rest of the four syllables were paired with neutral words such as "in", "of" or "up". After training, participants were provided with a six-page booklet. On each page, there were a seven - points semantic differential scale, which was a ranking scale about pleasantness and a nonsense syllable. Number "1" means pleasant and number "7" means unpleasant. Participants were required to rate the nonsense syllable from 1 to 7, and reported if they had seen the syllable before. In the experiment, researchers also asked participants' thoughts on those nonsense syllables as well as the purpose of the experiment to evaluate if the learnt equivalence relation (the meaningful word and nonsense syllables) requires participants' awareness of the purpose of the study. According to the data, most participants in group one rated YOF as around 2.4 and XEH around 4.73. For participants in group two, the average number for YOF was 4.8 and XEH was 3.13. It was shown that participants rated the nonsense syllables that were previously linked with the negative words lower - closer to 7, which indicated that the training was effective and had changed participants' preference towards certain nonsense syllables.

Tryron and Cicero (1989) reproduced Staats and Staats' (1957) study and extended their procedure to a third-order classical conditioning circumstance. It also aimed at exploring the boundary conditions linked with language conditioning. The stimuli used in the study were nonsense syllables and words, which were the same as those used by Staats and Staats (1957), except that there were two sets of nonsense syllables (CS2 and CS3), one paired with the word (CS1) and one paired with the other nonsense syllable (CS3 → C2 then CS2 → CS1). They wanted to see if the second set of nonsense syllables would take on the function of the words. After training that conditioning, participants were offered a seven-point rating scale, as in Staats and Staats' (1957). The results showed that the effect of language conditioning only transferred to CS2 but not to CS3. In the same year, Cicero and Tryon (1989) replicated Staats and Staats's research once more. The purpose and procedure of the research were the same as the previous one, except that they used "a triplet association" (CS3 → CS2 → CS1). In the training, the first nonsense syllables were presented visually on the screen with a projector for 1 s, which was followed by the second nonsense syllable presented for 1 s and the examiner simultaneously read the paired words loudly. The participants were asked to repeat the words immediately after the examiner finished pronouncing it. Again, according to the result, the conditioning effect did not show to the CS3.

Although there is mixed evidence for the effectiveness of higher-order

conditioning in humans and other animals, it is still a common assumption that it plays an important role in advertising (Pornpitakpan, 2012). Advertisements often associate a product and brand with other stimuli that are likely to elicit positive emotional responses. For example, the Coca-Cola® Company usually uses themes such as friendship, youth, and “fun” in their advertisements, which might increase the purchase or consumption of the product. It appears that classical conditioning is an important mechanism here, but there may be other mechanisms at work as well. The “dark side” of advertising is that when it comes into effective, the influences may cause people to buy things which might be harmful, such as cigarettes, alcohol, and soft drinks. Using cigarettes as an example, some tobacco advertisement associates smoking with the features of “success”, “independence” and “charm,” this may cause people to purchase and smoke more cigarettes. According to National Cancer Institute (2008), mass media advertising such as television commercials, billboards, and cinema take some of the responsibility for promoting the consumption of cigarettes. Moreover, National Cancer Institute (2018) suggested that advertisers can affect the population's social perceptions toward smoking by making use of some scenarios in movies or television shows, in which associated smoking with intimacy, social activity, and certain mood states such as "agitation, sadness, happiness, relaxation and pensiveness" (p. 373). According to Mothersbaugh and Hawkins (2015), what the consumer learned was not the information about the stimulus, but the emotional response to the stimulus, which will lead to consumers' engagement with the product. Set McDonald's “Dead Dad” advertisement as an example, a boy finds his father's stuff that have been left behind since he died, and he runs to the kitchen to ask his mother what his father looks like. His mother starts to recall all of the positive characteristics her husband has; then the mother and son enter McDonald's and the boy orders a hamburger combo. The mother smiles and tells the boy that this hamburger is his father's favourite food too; the boy and his father have the same table manners, with sauce dripping down his chin, and the mother looking at her son and smiling warmly. At the end of the advertisement, a big yellow letter M with a red background shows on the screen (anarchi.st, 2017). In this advertisement, the story of the dead father may function as a previously conditioned CS, and the emotions (i.e., physiological responses) related to memories of the father are the CR. The site of the hamburger (which may already serves as a CS) and the big yellow letter M showing at the end of the advertisement serve as the second-order CS. As a result, the big yellow letter M may come to elicit the emotions (CR).

Stimulus equivalence and advertising

Stimulus equivalence describes a set of procedures and outcomes from those procedures, which are commonly observed with humans. Equivalence is said to be established if three conditions are met: reflexivity, symmetry and transitivity (Barnes-Holmes et al., 2004). Reflexivity refers to an individual matching a stimulus to itself. For example, an individual matches a picture of a dog with another picture of the same dog. Symmetry means an individual can match two stimuli in reverse order without the reverse matching being trained. For example, after training an individual to match a real dog with a vocal response, “dog” without learning, an individual can match a vocal response, “dog” with a real dog. Transitivity describes the situation where A=B and B=C is trained and, without further training, A=C is demonstrated. For example, after an individual has learned to match a real dog with a vocal response, “dog” and a vocal response, “dog” with the written word, DOG, without further training, the individual can match a real dog with the written word, DOG. If all three are demonstrated, then the individual has demonstrated “equivalence”. When equivalence has been demonstrated, it has also been found that the function of the original stimulus can be transferred to an arbitrary novel stimulus, even in the absence of any direct training with the two stimuli. For example, if interacting with dogs functions as a reinforcer, then the word “dog” and DOG would inherit some of those reinforcing properties if equivalence is demonstrated.

Equivalence may be relevant, even critical, to advertising outcomes. Using Coca-Cola® as an example, advertisers associate enjoyable activities with cola drinking, and then build a relation between cola drinking and the brand. For instance, viewers may be shown images of smiling people on a hot day drinking cola and then be shown a glass of ice cold cola with the brand. The respondent (and reinforcing) functions already associated with smiling people and “thirst quenching” may be transferred indirectly to the brand and, as a consequence, the probability of buying and consuming Coca-Cola® may increase. Some researchers have attempted to evaluate the role of the processes responsible for stimulus equivalence in the transfer of reinforcing functions to neutral stimuli. Since the failures of respondent conditioning of transferring function to third-order conditioning, Barnes-Holmes et al. (2000) conducted three experiments using a stimulus equivalence procedure to examine if the function of emotive words could transfer to novel soft drink “brands”. The first experiment consisted of three stages: training, testing and rating. In training, the researchers used a matching-to-sample procedure to train the participants two types of relations. For example, when the word CANCER was shown, participants had

to select the nonsense syllable VEK. This nonsense syllable was later associated with "BRAND X"; when the word HOLIDAY was shown, participants had to select the nonsense syllable ZID. This nonsense syllable was later associated with a “BRAND Y” In the same way, participants were trained to choose ZID in the presence of HOLIDAY and choose BRAND Y in the presence of ZID. After that, a Matching to sample test was introduced. During the test, the brands (BRAND X and BRAND Y) were shown on the screen for 2 s, and then the screen cleared for 0.5 s. After that, two nonsense syllables were presented. Participants were instructed to select the one that in corresponding with the sample stimulus. Participants who got more than 17 correct out of 20 trials were treated as passed the test, whereas participants with less than 17 correct out of 20 trials were regarded as failed the test. In the last stage, every participant was offered a pleasantness rating scale. The score was from 1 to 7, which represented least pleasant and most pleasant respectively. Two bottles of cola were provided with one labeled “Brand X” and one labeled “Brand Y”. Participants were asked to taste the two bottles of cola and then rate both colas on the pleasantness scale. Most of the participants who passed the test rated BRAND Y (the positive brand) higher than BRAND X (the negative brand). The results showed that 27 participants passed the test and 9 participants failed. For those who passed, the conditioning appeared to influence their preference towards a certain brand. Even though the results supported the researchers' hypothesis, that a stimulus equivalence procedure would be successful in establishing third-order conditioning, it was unclear whether the successful transformation of function was due to the MTS procedure or the equivalence test (because the equivalence test was conducted prior to the preference test). Therefore, in order to answer this question, the equivalence test was removed from the second experiment. The second experiment recruited 8 new participants. The rest of the procedure remained the same. The outcome showed that 6 out of 8 showed preference for "holiday cola" and only 2 participants preferred “cancer cola”, indicating that the equivalence test was not indispensable in transferring preference. However, one possible criticism raised by the previous two experiments was that participants may have pre-existing preferences for the “brand” stimulus in the absence of “conditional discrimination training” (p. 504), therefore, a third experiment with the same participants was conducted. In this experiment, six new and experimentally naive participants were recruited for the study. Researchers first asked participants to rate colas as well as two words, HOLIDAY and CANCER before training. Next, training and testing stage were introduced, expect that this time the two words used for conditioning of BRAND X and BRAND Y were reversed (i.e., BRAND X was paired with CANCER and BRAND Y was paired with HOLIDAY). Then, rating for the pleasantness of the colas and the words CANCER and HOLIDAY

was conducted again by the participants. Comparing the pre and post pleasantness rating scores provided stronger evidence that "the conditioning discrimination training functioned as a powerful determinant of the subjects' preference rating" (p. 507). Four limitations were proposed by the researchers. First, the participants may have been aware of the experiment's purpose and, therefore, responded in a way that corresponded with the researchers' expectations. Second, as the two brands of cola were presented to participants at the same time, as a consequence, the evaluation of participants toward one brand may have transferred to the other brand. Third, comparing the result among the three experiments, more participants preferred the "holiday" brand in Experiments 2 and 3 than in Experiment 1. It is paradoxical that without receiving equivalence test, participants showed "greater transfer of emotive functions" (p. 509). The authors suspected that the repetitive exposure to the words (Cancer and Holiday) in Experiment 1 may increase any habituation of participants toward the emotive functions of the words. Fourth, the perception of participants to the words CANCER or HOLIDAY (i.e., the function of these words) may be different from the perception of researchers. The researchers may have treated the word CANCER as negative and HOLIDAY as positive according to their own personal histories. However, some of the participants may perceive the word CANCER as positive and the word HOLIDAY as negative due to their personal history (e.g., a participant might have had a tragic accident on a holiday).

Smeets and Barnes-Holmes (2003) conducted two experiments looking at the preference for novel soft drink brands in children by using a stimulus-equivalence procedure similar to the one used in the previous study. The first experiment consisted of three parts: matching to sample training, equivalence test, and preference assessment. A matching-to-sample task was conducted to teach relations among six stimuli. These stimuli were a smiling cartoon character (A1), a picture of a crying baby (A2), two black geometric forms (square - B1, triangle - B2) and two arbitrary symbols (C1 and C2). There were 4 conditions. Conditions 1 and 2 used forward conditioning. In condition 1, participants were trained to select A1 with the presence of B1, and select A2 with the presence of B2, and select B1 with the presence of C1, and select B2 with the presence of C2. Condition 2 was the same, except that this time participants were asked to select B2 in the presence with C1 and select B1 in the presence of C2. Condition 3 and 4 used backwards conditioning. In Condition 3, participants were trained to select B1 in the presence of A1 and select B2 in the presence of A2 and select C1 in the presence of B1 and select C2 in the presence of B2. The experimenters reversed the C-B relation in condition 4. Participants' correct responses resulted in the reward of a bead. The beads were used to fill a glass container. Once the bead reached a mark on the container, the

experimenter would empty the container and let the participants choose one preferred picture card and continue the training. After the training, in order to assess whether participants demonstrate transitivity, which would involve matching A stimuli with C stimuli, an equivalence test was introduced. The participants were offered a second chance if they failed the test. During preference assessment, two cups of lemonade with labels C1 and C2 were placed on the table, and participants were required to choose the one that they preferred to taste first. After the data were collected by the experimenter, the participants were asked to taste both cups of soft drinks and tell the experimenter which one they thought tasted better. Next, the experimenter presented the two A-stimuli (cartoon character - A1 and crying child - A2) to the participants and let them choose which one they liked best. For the purpose of determining if the equivalence test is essential for the occurrence of transfer, a second experiment was conducted with the equivalence test removed. The results from both experiments showed that after having tasted both drinks, 75% of participants preferred the drink which was associated with the cartoon character in Experiment 1 and 88% participants preferred the drink which was associated with the cartoon character in Experiment 2, which indicated that participants were strongly influenced by the matching-to-sample task regardless of the equivalence test. However, there were several limitations associated with this study. Firstly, the validity of the preference test may be doubted since two participants failed the equivalence test but showed transfer. Furthermore, the study did not answer the question about why the respondent conditioning procedure used by Cicero and Tryron (1989) did not produce third-order conditioning but the present study appeared to do so. Such dissimilarity may be due to the different number of the stimuli that had been used. Cicero and Tryron (1989) used "108 different CS2 → CS1, CS3 → CS2 sequences so that no stable associations could be formed between the nonsense syllables and any given word" (Cicero & Tryron, 1989, p. 139). Additionally, the methods were quite different in that Tryon and Cicero (1989) used a respondent conditioning procedure whereas Smeets and Barnes-Holmes (2003) used an operant procedure. However, the use of an operant procedure represents another limitation of the present study, because operant procedures do not seem to be appropriate for use in studying mechanisms at work in advertising because individuals are rarely required to interact with an advertisement. Moreover, as Smeets and Barnes-Holmes (2003) mentioned, the conditioning effect is likely to be short-lived since participants had the test immediately after the conditioning training. Lastly, since people normally see advertising under less constrained situations, highly controlled settings may influence the external validity of the study.

Valdivia-Salas et al. (2013) conducted three tests to investigate the influence of

evaluative conditioning on alteration of preference and its generalizability. Participants were distributed in three groups with three different conditions – Group 1 received evaluative conditioning and equivalence training, Group 2 received evaluative conditioning but not equivalence training, and Group 3 received neither evaluative conditioning nor equivalence training. The experiments first established four geometrical shapes (square, circle, hexagon and triangle) as discriminative stimuli for selecting four pictograms, B1, B2, C1 and C2, respectively. Then, the experiment evaluated B1 and B2, C1 and C2 by using a simultaneous discrimination task in which participants selected one of two geometrical shapes that were shown on the screen and, immediately following their selection, were shown the corresponding pictogram. In the baseline assessment, B1 was paired with unpleasant slide and noise, and B2 was paired with a pleasant slide (for Groups 1 and 2). The unpleasant slide presented damaged and/or disfigured bodies, and the pleasant slide presented some funny scene with babies and/or pets. Then, the experiment equated B1 to C1 and B2 to C2 using equivalence training and testing. According to the results, the majority of participants in the experimental groups chose B2 and C2 compared to the control group. Afterwards, the experimenters exposed all participants to equivalence training to establish the geometric shapes as equivalent to the letters X, Y, W and Z, respectively. Then, the letters were presented first in the simultaneous discrimination tasks. As expected, only participants in the experimental conditions preferred letters corresponding to the shapes that were equated to B2 and C2. One concern mentioned by authors is that, since the experiment was conducted in a controlled setting, it remains unclear if in an uncontrolled setting, the same outcome would be obtained. According to the article, most participants demonstrated third- and even fourth-order conditioning, which showed that the authors successfully established higher order conditioning by using the equivalence procedure.

Respondent vs. Operant Procedures

By far, matching to sample (MTS) is the most frequently used format for equivalence training. However, as mentioned, this procedure does not have good external validity when it comes to advertising because people do not usually have to interact with advertisements, yet they still seem to be effective. Kinloch et al. (2013) compared the effectiveness of a respondent approach to equivalence training, the Stimulus Pairing Observation Procedure (SPOP) with the standard operant approach (MTS). Ninety-four participants were recruited for the study and were equally divided into twelve groups. These groups were different in "the procedure (SPOP or MTS), the arrangement of the stimuli (linear, many-to-one, or one-to-many) and number of

training trials completed (60 or 120) (p. 161)". There were nine stimuli (CUG, ZID, VEK, YIM, PAF, ROG, MAU, JOM, and DAX) which were equally divided into three groups (A1, A2, A3; B1, B2, B3; C1, C2, C3). Each participant in each session needed to complete the training and testing cycles twice. There were eight sessions in total. During SPOP, participants were presented with pairs of stimuli. The sequence was as follows: the first stimulus of each pair shown on the screen for 1 s, which was followed by a 0.5-s blank, and then the second stimulus of each pair appeared for 1 s, which was followed by a 3-s between-pair delay. Each pair was presented 10 times for the participants in 60 training trials or 20 times for the participants in 120 training trials. The procedure for MTS procedure was as follows: one sample stimulus was shown on the top center of the screen, three comparison stimuli later were shown on the bottom of the screen, participants were asked to select the correct stimulus on the bottom to correspond to the sample stimulus. A 1000 Hz tone and the word CORRECT would show if the participant did a correct response and the word INCORRECT would appear if the participant did an incorrect response. After the training (SPOP and MTS procedure), a matching to sample test was introduced for the purpose of testing if participants generate derived relations. The test procedure was similar to the MTS procedure, except that participants who failed the test (below 9 out of 10 correct responses for each of the nine stimuli) were required to do the training again. According to the result, overall, for SPOP and MTS procedure, the percentage correct was 61.7% and 68.1%, respectively; which indicated that SPOP was nearly as effective as the operant method. For linear, many-to-one or one-to-many stimulus arrangement, the percentage of correctness was 43.8%, 68.8% and 83.3% respectively. For overall 60 training trials per cycle and overall 120 training trials per cycle, the percentage of correctness was 55.9% and 70% respectively. Moreover, compared to the MTS method, SPOP has better external validity because this method is more similar to real-world marketing methods. MTS procedure requires subjects' overt selection response, whereas, for SPOP, the subjects are only required to observe the stimuli being presented together (Clayton & Hayes, 2004).

Imagery for Conditioning and Advertising

Though words are used in advertising, advertisers often employ imagery in addition to words or exclusively use imagery. Smeets and Barnes-Holmes (2003) used a smiling cartoon character and a picture of a crying baby as the positive stimulus and negative stimulus, respectively. In some anti-smoking campaigns, negative images are used with the intention of influencing people away from smoking. Mead et al. (2015) examined whether graphic warning labels on cigarette

packaging could change individual's smoking-related behaviour. In the research, they used positive and negative pictures to demonstrate the benefits of quitting smoking and the harmfulness of smoking, respectively. Participants were asked to give a self-report related to their "cognitive and affective reactions to each label" (p. 3). The results suggested that graphical warning labels on cigarette packaging can induce emotional reactions in line with smoking cessation, especially for the pictures that show the negative consequence of using cigarettes.

The Current Study

The current study was aimed at evaluating the effectiveness of second-order and third-order conditioning using stimulus equivalence procedures with the goal of improving the internal and external validity of the research in this area. A respondent procedure (SPOP) was used instead of an operant procedure, such as MTS, because the former seems to have better external validity for research in this area. The present study employed imagery, rather than words, in the conditioning procedure because this also seems to have better external validity. Additionally, because none of the previous studies included a neutral condition (i.e., where the “emotive” CS is unlikely to have any significant behavioural functions), it is unclear if participants in those studies were choosing away from the negative stimuli or towards the positive stimuli or a combination the two. Therefore, in the present study, a neutral condition was included as well. In line with the previous studies by Smeets and Barnes-Holmes (2003), the influences of equivalence training on participants’ preference for “brands” of soft drink was also evaluated in this study.

Method

Subjects

Twenty participants, twelve female and eight male, participated in the research. Their age was between 15 and 30. Participants were recruited from the university and the surrounding area using advertisements or through personal contact. All participants had no previous experiences with the SPOP or the MTS procedure.

Before the experiment began, consent was obtained. Each of the participants was told that some negative pictures such as rotten teeth, lost toes and heart disease surgery would be used during the research which might cause them to feel distressed.

The participants were allowed to withdraw from the study at any point and were provided with contact information for Student Counselling Services and told that they could receive support from them if required. Ethical approval was obtained from the University's Psychology Research and Ethics Committee (17:54).

Procedure

The experiment consisted of three parts: training, preference assessment, and matching-to-sample (MTS) test. Training was conducted in two sessions on separate days within a two-week period. Immediately following the second training session, two preference assessments were conducted, one MSWO procedure and one taste-test ranking procedure. During the taste-test ranking procedure, participants were asked to taste and use a scale to rank the degree of preference for the six brands. Next, a matching to sample test was conducted to evaluate the effectiveness of the training procedure. Each of the two experimental sessions took approximately 30-40 minutes.

Training

In this stage, nine pictures with different anticipated respondent functions were used, three negative pictures, three positive pictures, and three neutral pictures. Each group of pictures was paired with a brand either directly or indirectly (through another arbitrary symbol) and, therefore, there were six brands. Symbols that did not appear to have any relationship to commonly used symbols were selected by the researcher in Microsoft Word. The negative pictures that were chosen were related to typical outcomes of excessive sugar consumption, such as heart disease, tooth decay and diabetes. The positive pictures that were chosen were related to health and fitness levels associated with good dietary choices. Similar negative pictures had been used in anti-tobacco marketing campaigns in New Zealand and many other countries (see Appendix 1 for a selection of pictures and symbols that were chosen for use in this research). Before selecting the specific pictures used in the study, the researcher asked eleven people to choose the pictures that "produced the biggest positive and negative emotional reaction" and selected those with the highest rating.

For indirect conditions, participants were exposed to associations between the pictures and an arbitrary symbol. For the direct conditions, the negative, positive, and neutral pictures were related directly to a "negative", "positive", or "neutral" arbitrary symbol, the "brand" of soft drink. During SPOP training, a total of 18 stimuli - nine pictures (three of each type) and nine arbitrary symbols (six brands and three

“bridging stimuli” for the indirect pairings) - were presented to the participants in the form of 21 stimulus pairs (See Appendix 2). Each stimulus pairing occurred in the center of the computer screen. The study employed a standard forward-pairing respondent conditioning procedure in which the neutral stimulus (e.g., brand) was displayed on the screen for 1 s and immediately followed by the conditioned stimulus (e.g., picture). Between each pairing, the screen was blank for 2 s. After three randomly selected stimulus pairings, the participants were required to complete a short memory test. The memory test was arranged as follows: Eight stimuli (four pairs of stimuli) were shown on the computer screen (See Appendix 3), and the participants were instructed to choose the stimulus pair that they did not see in the previous three stimuli pairings. A Green Check appeared on the screen for 1 s following a correct choice and a Red Cross following an incorrect choice. If the participant chose the correct pair, the training continued with the next stimulus pairings. If they chose incorrectly, the participant was shown the same three pairings again and given a second memory test. However, if the participant chose incorrectly a second time, the training resumed with the next stimulus pairs. Timing between the last pair and the memory test was 1 s, and the time between the memory test and the first stimulus presentation was 2 s. There were 210 trials in each session. After every 51 trials, and the final 57 trials (instead of 51 trials), the screen displayed a message with the percentage of memory tests that the participants got correct (excluding the tests following retraining trials). The statement read "So far you have gotten X% correct!" and at the end of the session, it read, "You got X% correct" on the last feedback screen.

MSWO and Ranking

After training and testing, participants received a preference assessment. The preference assessment had two parts: Multiple Stimulus Without Replacement (MSWO) and taste-test ranking. In the MSWO assessment, 6 bottles (300ml) filled with cola with a picture of each of the brands (2cm by 3cm) were placed in a line on the table within the same distance of each other and the participant. Other than the different brands, the bottles were identical. The experimenter asked participants to choose one of the six bottles of soft drink that he/she most wanted to drink by saying “Please point to the one you prefer.” Once the symbol of soft drink was selected by the participant, it would not be replaced in the next trial. Before the next trial started, the soft drink at the left end of the line was moved to the right end, and the rest of the soft drinks were also shuffled around. The procedure for the rest of the trials remained the same until only one remained. Each participant completed the MSWO twice. The researcher ranked the symbol (brand) from 1 to 6 according to the

sequence it was selected by the participants. The lower the number, the higher the degree of preference (See Appendix 4). Following the MSWO assessment, each participant tasted 6 glasses of soft drink out of new disposable cups with the 6 arbitrary "brand" symbols attached to the cups. The soft drink in each cup was the same, a diet (sugar-free) cola. Before the participant finished the second session of training, the researcher left the experimental room and went to another separate room about a hundred meters away to prepare the cups with cola and cups with soda water, and then used a tray to bring them to the participants. The participants did not have access to the room and were not told that all of the cups were filled from the same bottle of cola. The capacity in each cup was 30ml, only 10ml cola was offered to participants. A ranking arrow with BETTER printed at one end and WORSE printed at the other was placed on the table in front of the participants (See Appendix 5). The researcher pointed to the cups and the ranking picture and said: "Please taste the drinks in these cups one by one and place the cup on the line depending on if it tasted better or worse than the other one." There was a cup of soda water beside each soft drink for the participant to refresh his/her mouth, and they were instructed that they could do so. Participants' placement of the cups on the scale was used to rank the 6 soft drinks from 1 to 6. The smaller the number, the higher the degree of preference.

Matching to Sample Test

After the participants had completed the training and the Preference Assessment, in order to examine whether indirect pairing was effective in establishing a relationship between the indirectly paired stimuli, each participant completed a Matching to Sample test. In the MTS, three pictures and one arbitrary symbol were presented on the computer screen, with the picture on the top and three arbitrary symbols on the bottom (See Appendix 6). The top picture appeared first, then the three options were presented at the bottom after a 1.5 s delay, participants were required to choose one arbitrary symbol (brand) from the bottom corresponding to the picture on the top. There were approximately 90 trials in total. The table (See Appendix 7) shows an example of the matching-to-sample procedure. The table described the stimulus names, which were exactly what were used in the SPOP programme and described the stimuli that were presented in each MTS trial. Each of the trials was run twice in a randomized order. The arrangement of the three "brand" stimuli on each trial was also be randomized. All direct and indirect (derived) pairings were tested and the "bridging stimuli" were not presented in the MTS procedure.

Results

Figure 1 displays the mean scores on the memory tests from the first training session and the second training session for each participant and the mean across all participants. Mean scores on the two memory tests remained above 80% for all participants with the exception of Participant 13, whose mean score was 70.93% in Session 1. In general, the fluctuation of the memory test score in Session 1 was greater than in Session 2. Overall, the accuracy of participants in the memory tests was similar in Session 1 and Session 2.

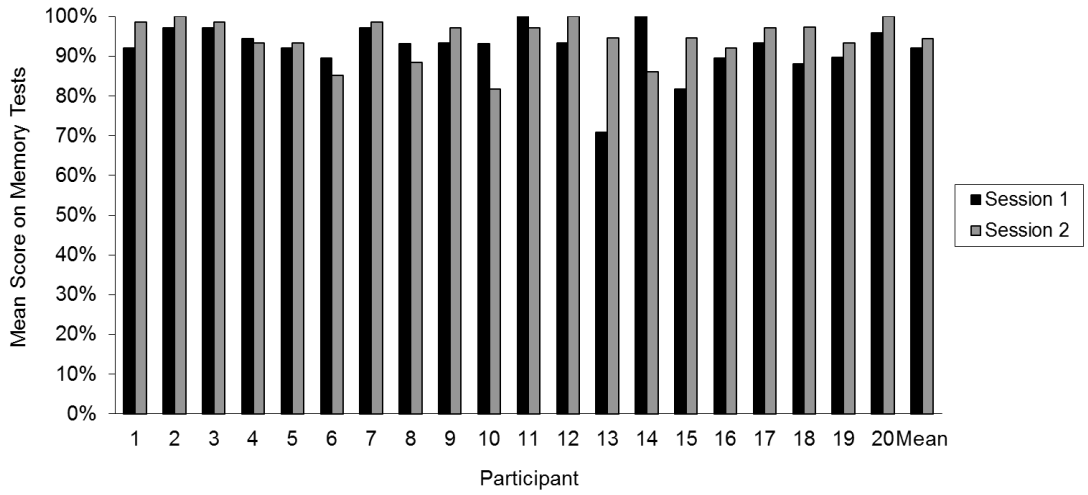


Figure 1. Mean scores on the memory tests from the first training session and the second training session for each participant and the mean across all participants.

Table 1 displays the mean, standard deviation, and 95% CI for the MSWO preference assessment results. The average score for all participants was near 3.5, except for the Direct-Negative brand, which was reliably selected close to last (4.75) and for Indirect-Neutral brand, which was selected earlier than other brands, on average (2.48).

Table 1.

Mean, standard deviation, and 95% CI for MSWO preference assessment data.

		Mean	SD	95% CI
Direct	Negative	4.75	1.464	4.065, 5.435
	Neutral	3.50	1.662	2.722, 4.278
	Positive	3.475	1.666	2.695, 4.255
Indirect	Negative	3.475	1.352	2.842, 4.108
	Neutral	2.475	1.602	1.725, 3.225
	Positive	3.375	1.385	2.727, 4.023

Figure 2 displays the mean order each brand was selected in the MSWO assessment across all participants. Both negative and neutral directly paired brands were selected later (i.e., lower preference) compared to the corresponding indirectly paired brands, but for positive brands, there was little difference in the mean order of selection between directly and indirectly paired brands.

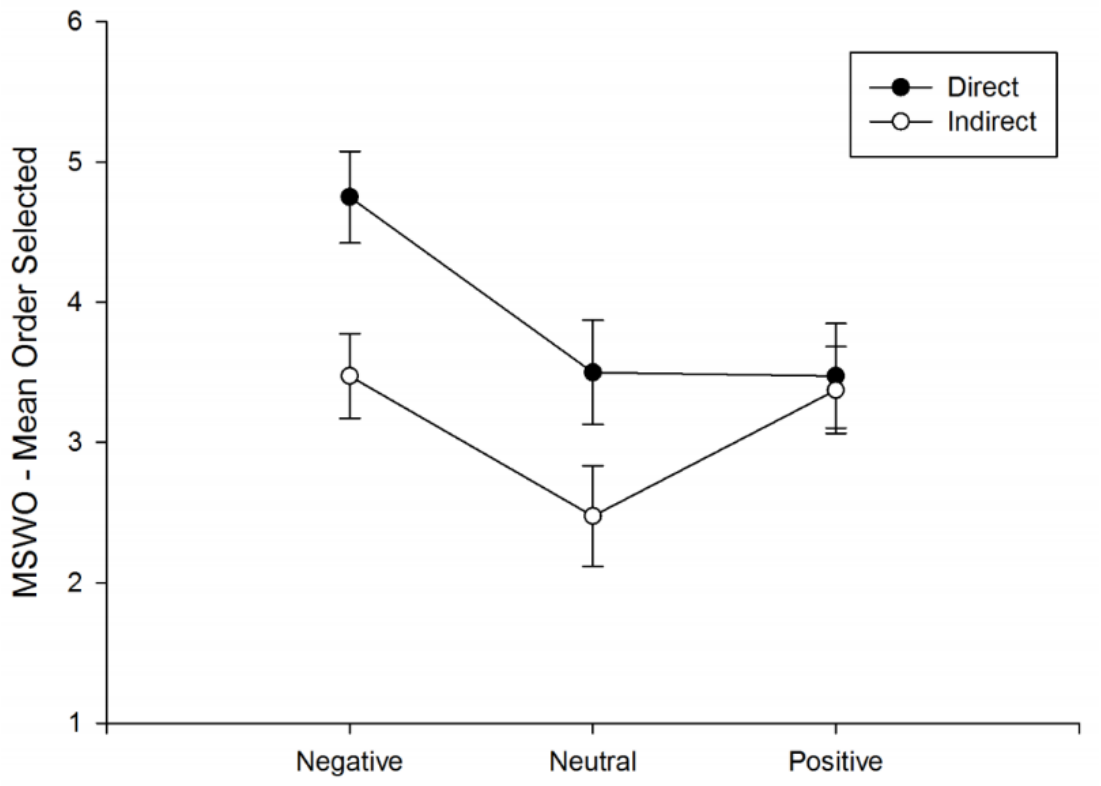


Figure 2. Interaction plot displaying the mean order each negative, neutral, and positive brand was selected for both directly and indirectly paired brands in MSWO trials for all participants (with SE bars). Note that a lower score indicates that a brand is selected earlier (i.e., higher preference).

A 3x2 repeated-measures ANOVA with the training type (direct or indirect) and the picture type (negative, neutral or positive) as Within-subjects factors for MSWO revealed main effects of training type, $F(1, 19) = 5.302$, $p = .033$, $\eta_p^2 = .218$, and picture type, $F(2, 38) = 4.38$, $p = .002$, $\eta_p^2 = 0.277$. These main effects were not qualified by an interaction between the training and picture type, $F(2, 38) = 1.121$, $p = .336$, $\eta_p^2 = .056$.

Table 2 displays the mean, standard deviation, and 95% CI for the taste-test ranking preference assessment. The mean ranking for all brands was near three, the middle position, except the mean for the Direct-Negative brand, which was 3.95, which is slightly higher than other conditions (i.e., lower preference).

Table 2.

Mean, standard deviation, and 95% CI for RANK preference assessment data.

		Mean	SD	95% CI
Direct	Negative	3.950	1.669	3.169, 4.731
	Neutral	3.750	1.888	2.866, 4.634
	Positive	3.300	1.720	2.495, 4.105
Indirect	Negative	3.050	1.504	2.346, 3.754
	Neutral	3.500	1.821	2.648, 4.352
	Positive	3.450	1.731	2.640, 4.260

Figure 3 shows the brands paired directly and indirectly plotted across the picture type. It can be seen that participants' preference was neither affected by the picture type, nor the training type. Except, there appears to be a difference in directly and indirectly paired negative brands.

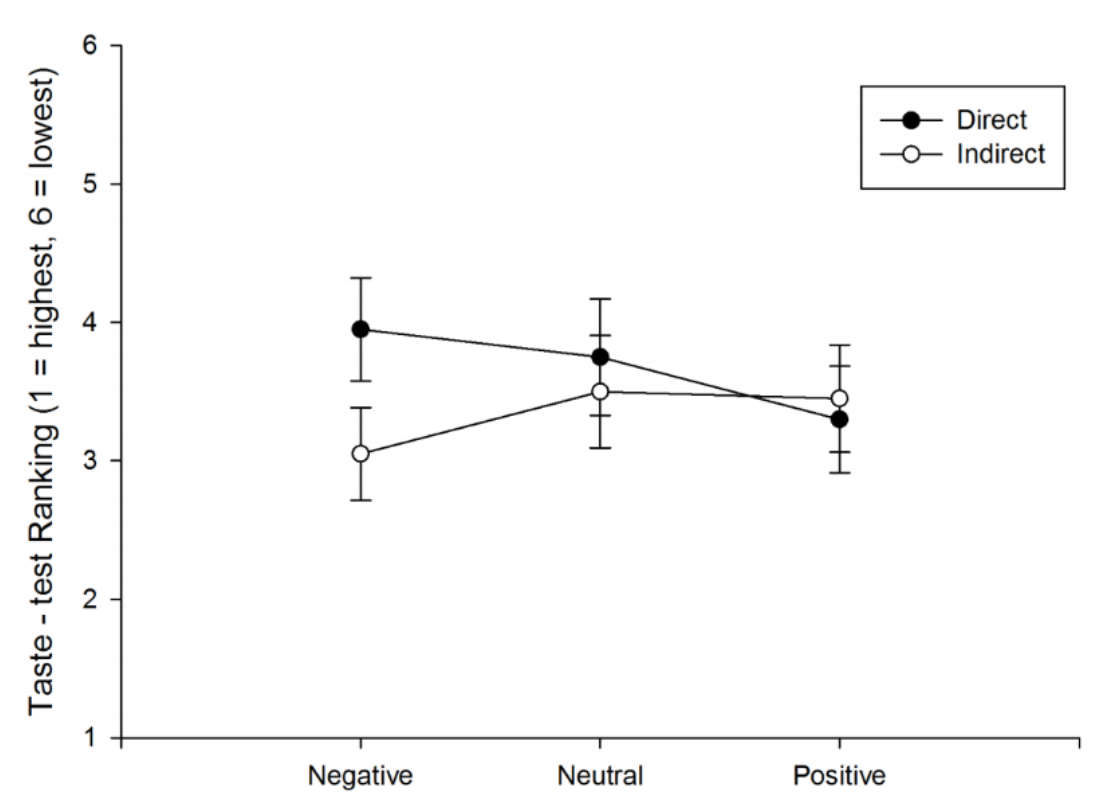


Figure 3. Interaction plot displaying the mean taste-test ranking for each negative, neutral, and positive brand for both directly and indirectly paired brands for all participants (with SE bars). Note that a lower score indicates that a brand is ranked higher (i.e., higher preference).

A 3x2 repeated-measures ANOVA with the training type (direct or indirect) and pictures type (negative, neutral or positive) as Within-subjects factors for Rank revealed that these are not statistically significant, $F(1, 19) = 0.633$, $p = .436$, η_p^2

= .032, and picture type, $F(2, 38) = 0.225$, $p = .800$, $\eta_p^2 = 0.012$. These is not statistically significant main effect for the interaction between the training type and picture type, $F(2, 38) = 0.796$, $p = .458$, $\eta_p^2 = .040$.

Participants' performance with each brand type in the MTS test is portrayed in Figure 4. It can be seen that the participants were more accurate selecting brands that matched directly paired pictures compared with indirectly paired pictures.

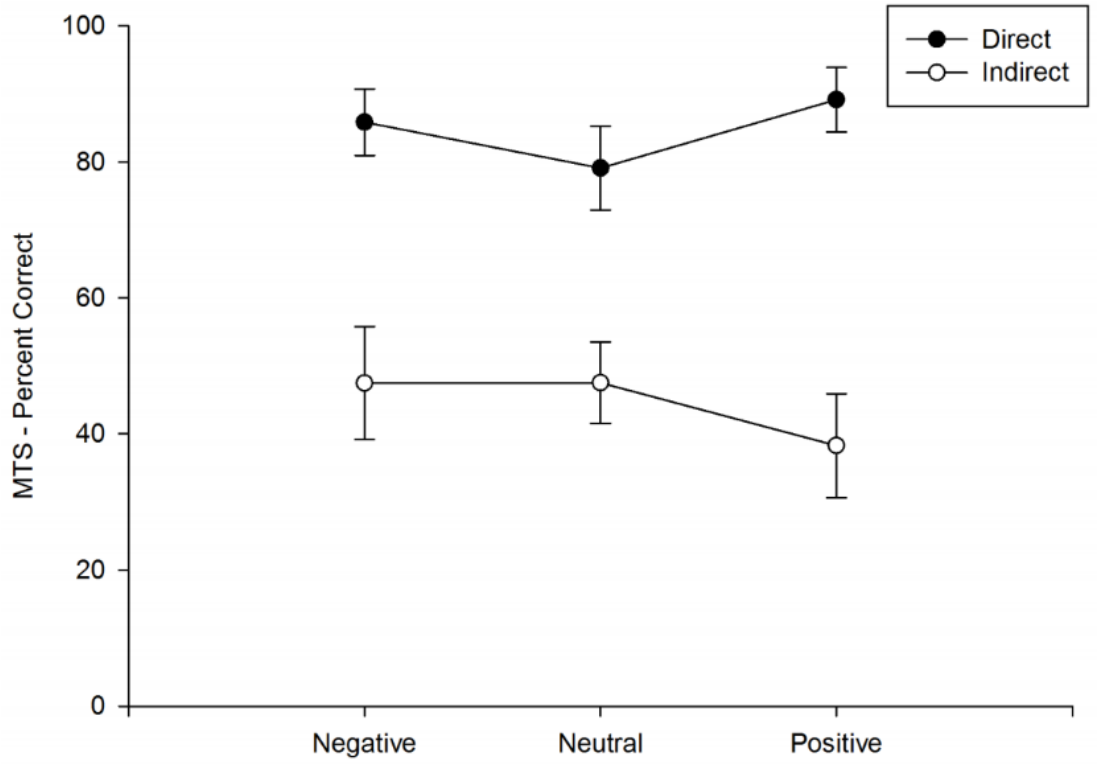


Figure 4. Interaction plot displaying the mean percentage correct in the MTS test for each negative, neutral, and positive brand for both directly and indirectly paired brands across all participants (with SE bars).

Accuracy data from the MTS test results were analysed with a 3x2 repeated-measures ANOVA to evaluate the influence of training and picture type on participants' accuracy. The Within-subjects factors revealed a main effect of the training type, $F(1, 19) = 66.395$, $p = .000$, $\eta_p^2 = .778$, but not picture type, $F(2, 38) = 0.188$, $p = .830$, $\eta_p^2 = 0.010$. These main effects were not qualified by an interaction between the training and picture type, $F(2, 38) = 1.307$, $p = .283$, $\eta_p^2 = .064$.

A Pearson correlation coefficient was calculated to examine if performance on the memory test correlated with performance on the MTS test. Figure 5 shows the relationship between the accuracy of memory test and MTS test. The small positive

correlation was not statistically significant, Pearson's $r(20) = .20$, $p = .394 > .05$

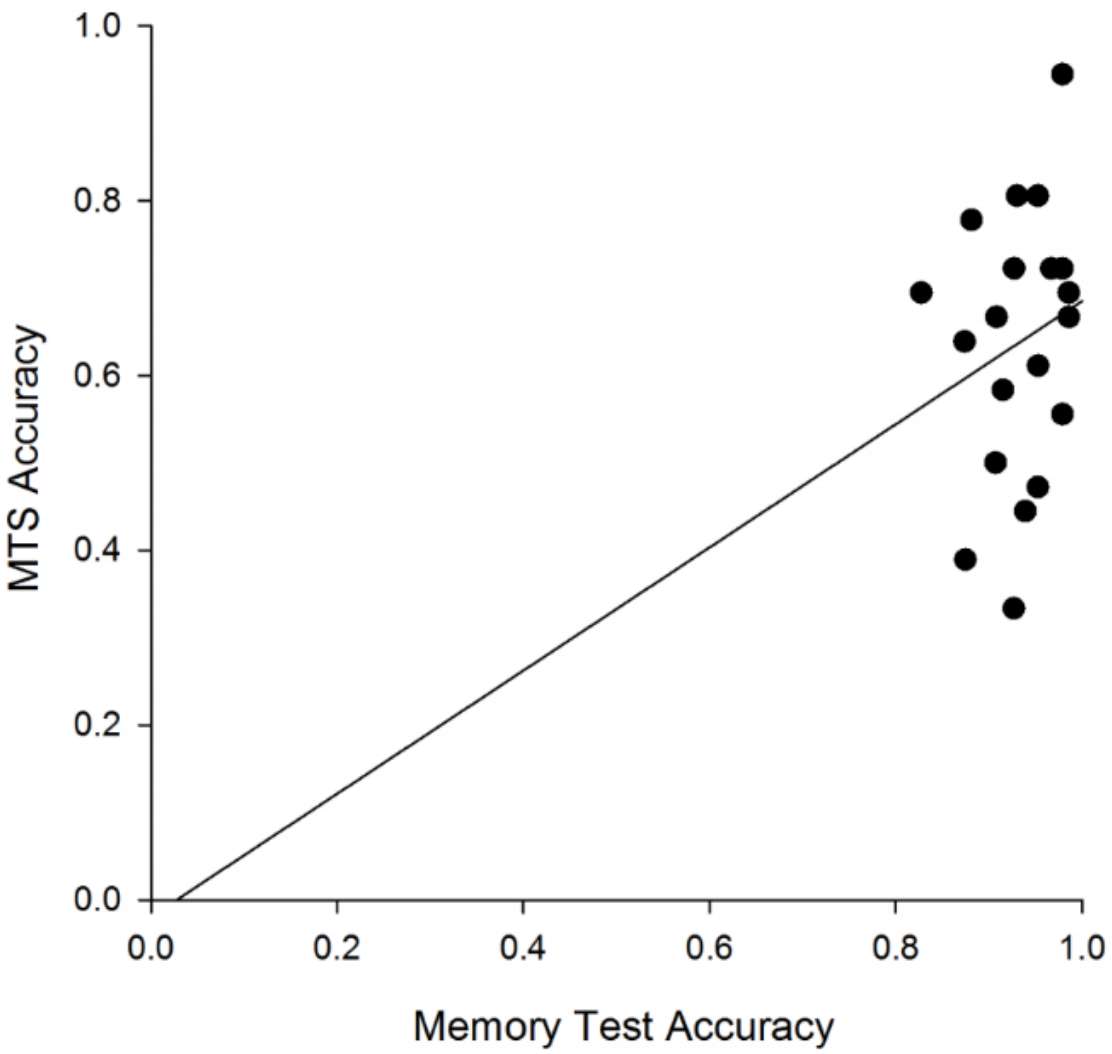


Figure 5. MTS accuracy plotted against Memory test accuracy.

Additional analyses of data from the MSWO and Ranking tests were conducted with participants who performed poorly (below 60% accuracy) in the MTS test excluded. Poor performance in the MTS test suggests that conditioning was not complete or effective for these participants and, therefore, would not be expected to impact their preference results as much as those participants who performed well in the MTS tests. A similar approach to analysis was adopted by Barnes-Holmes et al. (2000). Mean, standard deviation, and 95% CI of MSWO for high performers is presented in Table 3. Except for Direct-Negative and Indirect-Neutral brands, with mean order selected of 4.923 and 2.154 respectively, the mean order for other brands is around 3.

Table 3.

Mean, standard deviation, and 95% CI for MSWO-High preference assessment data.

		Mean	SD	95% CI
Direct	Negative	4.923	1.669	3.915, 5.932
	Neutral	3.577	1.412	2.724, 4.430
	Positive	3.192	1.614	2.217, 4.168
Indirect	Negative	3.770	1.166	3.065, 4.474
	Neutral	2.154	1.214	1.420, 2.888
	Positive	3.385	1.431	2.520, 4.249

Figure 6 shows the selection order of the directly and indirectly paired brands plotted across picture type. Overall, Direct-Negative and Direct-Neutral brands were selected later than their indirectly paired counterparts, but Direct-Positive was selected sooner than the Indirect-positive, on average. Interestingly, the Indirect-Neutral brand was selected earlier than both Direct-Positive and Indirect-Positive brands, on average.

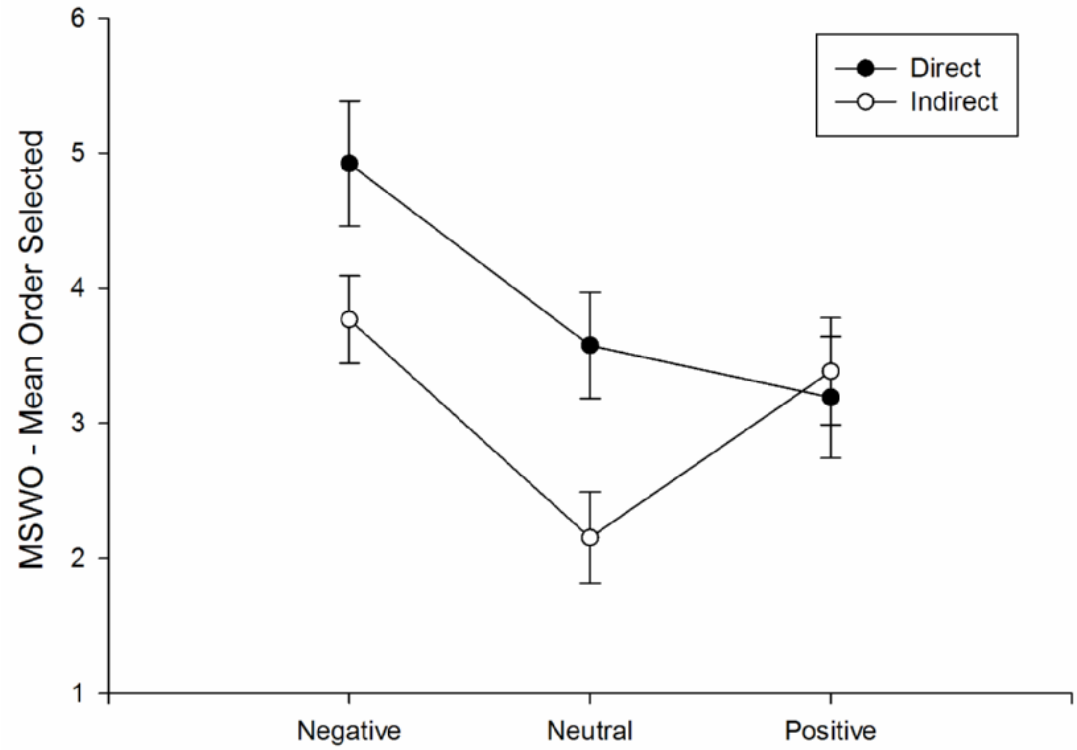


Figure 6. Interaction plot displaying the mean order each negative, neutral, and positive brand was selected for both directly and indirectly paired brands in MSWO trials for high performers (with SE bars). Note that a lower score indicates

that a brand is selected earlier (i.e., higher preference).

A 3x2 repeated-measures ANOVA with the training type (direct or indirect) and pictures type (negative, neutral or positive) as Within-subjects factors for MSWO (high performer) revealed a main effects of the training type, $F(1, 12) = 5.010$, $p = .045$, $\eta_p^2 = .295$, and picture type, $F(2, 24) = 8.541$, $p = .002$, $\eta_p^2 = 0.416$. These main effects were not qualified by an interaction between the training type and picture type, $F(2, 24) = 1.554$, $p = .232$, $\eta_p^2 = .115$. In order to clarify which means were significantly different from each other among negative, neutral, and positive brands, post hoc t-tests using Bonferroni corrections were conducted. The result indicated that the mean score of the negative brands was significantly different from the mean score of the neutral brands ($M = 1.481$, $SD = 0.437$, $p = .005$) and was significantly differ from the positive brands as well ($M = 1.058$, $SD = 0.356$, $p = .012$).

Table 4 shows the mean, standard deviation, and 95% CI of the taste-test ranking for high performers. It shows that the Direct–Negative brand and the Direct–Neutral brand were ranked the same - 4.15 on average. While the average ranking of all other brands was near 3.

Table 4.

Mean, standard deviation, and 95% CI for high performers' taste-test ranking preference assessment data.

		Mean	SD	95% CI
Direct	Negative	4.154	1.573	3.203, 5.104
	Neutral	4.154	1.725	3.112, 5.196
	Positive	3.154	1.819	2.055, 4.253
Indirect	Negative	3.077	1.605	2.107, 4.047
	Neutral	2.923	1.605	1.953, 3.893
	Positive	3.539	1.854	2.418, 4.659

Figure 7 displays the rankings associated with brands from direct and indirect training plotted across picture type (negative, neutral, positive). Direct-Negative and Direct-Neutral brands were ranked lower than their indirectly paired counterparts, but Direct-Positive was ranked higher than the Indirect-Positive.

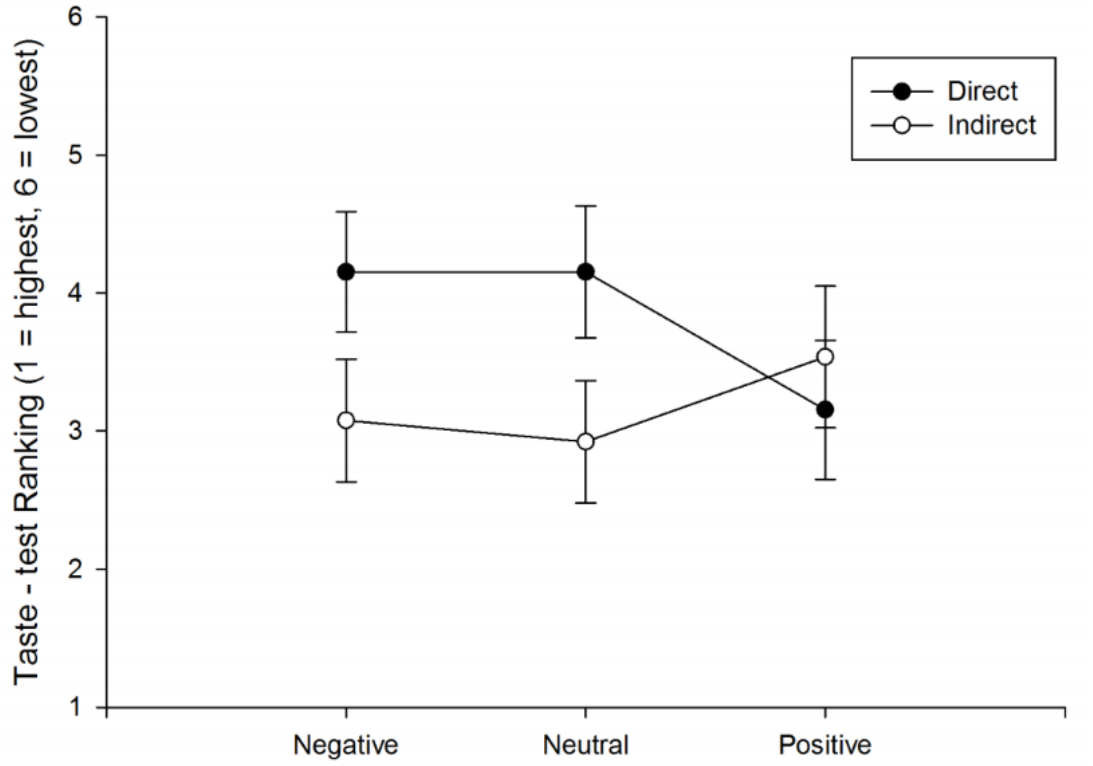


Figure 7. Interaction plot displaying the mean taste-test ranking for each negative, neutral, and positive brand for both directly and indirectly paired brands for high performers (with SE bars). Note that a lower score indicates that a brand is ranked higher (i.e., higher preference).

A 3x2 repeated-measures ANOVA with the training type (direct or indirect) and picture type (negative, neutral or positive) as within-subjects factors for Rank (high performer) revealed that the main effects for training type, $F(1, 12) = 1.210$, $p = .293$, $\eta_p^2 = .092$, and picture type, $F(2, 24) = 0.164$, $p = .849$, $\eta_p^2 = 0.014$ are not statistically significant. The interaction between the training type and picture type was also not statistically significant, $F(2, 24) = 2.228$, $p = .130$, $\eta_p^2 = .157$.

Discussion

The present study aimed at investigating the effectiveness between the indirect and direct pairings of pleasant and unpleasant pictures with product brands. During the experiment, participants were exposed to respondent training in which “pleasant”, “neutral”, and “unpleasant” stimuli were paired directly with arbitrary symbols, or indirectly through another arbitrary symbol. Unlike previous research, the training procedure in the current experiment used SPOP, a passive procedure, which is closer to real marketing methods than MTS, an active method.

Considering the preference assessment results from the participants who performed well in the equivalence/MTS test, directly paired brands were selected later in the MSWO test, overall (i.e., they were lower preference) and both directly and indirectly paired negative brands were selected later than their neutral and positive counterparts. These findings suggest that negative pictures were more effective than neutral and positive pictures at influencing preference; and the direct pairing was more effective than indirect ones in shifting preference. Descriptively, the taste-test ranking data from high performers corresponded with the MSWO data, but the effects of brand type were not statistically significant. This result was similar to those had been reported by Smeets and Barnes-holmes (2003), but the difference was that they used the MTS training to show that the participants' preference can be influenced by positive and negative pictures whereas the present study showed that similar results can be obtained with SPOP procedure.

Additional information provided by the current study was that, relative to "neutral" pictures, "unpleasant" pictures shifted preference away from associated brands more than "pleasant" pictures shifted preference toward associated brands. This phenomenon may be a result of individuals' exposure to many positive pictures in everyday life, including advertising, which may reduce the effectiveness of these pictures for respondent conditioning compared to the type of negative pictures used in this study, which are not normally presented in public. Therefore, letting participants rate positive and negative pictures prior to using them for research may be a good idea, but it does not resolve this potential issue of the individual's history with different amounts of exposure to positive and negative pictures.

Moreover, the study also provides additional information about the relative effectiveness of direct versus indirect conditioning. According to the result, regardless of whether the picture was negative, neutral or positive, participants performed better on MTS test when they learnt the relation between symbol and picture in a direct way. Accurate selection in the MTS task with direct brands was only a demonstration of reflexivity because the stimuli had been paired directly, but in the opposite order to the order presented in the MTS task, whereas accurate selection with indirect brands was a demonstration of equivalence because the stimuli in the MTS task had not been paired directly and the indirect pairings had been done in the reverse order. Therefore, this outcome is not surprising from a stimulus equivalence perspective or a respondent-conditioning perspective.

The current study included the memory test, which Kinloch et al. didn’t introduce in their study. The memory test aimed to help participants better attend to the stimulus pairings by introducing a consequence for doing so or failing to do so.

In this sense, the SPOP as conducted in this study did have some properties of an operant task. The correlation analysis between the MTS and memory test performance suggested that performance on the memory test was not correlated with their performance on the MTS test. Memory test scores were high across all participants and, therefore, there may have been a ceiling effect which may be the reason that no correlation was observed. For example, Participant 15 achieved an average of 88.15% accuracy on the memory test and 77.18% on the MTS test, which placed him in the high-performance group, whereas Participants 16 also got a high score on the memory test, 97.74%, but only obtained 50% accuracy on the MTS test.

The recall/remembering skills of the participants may also play an important role in the obtained results. More specifically, some participants may be better at this type of “recall” task (memory test) for a variety of reasons, and this may have impacted their ability to be conditioned. According to Hughes et al. (2016), in their experiment, they did contingency memory test with each participant to explore whether participants "could recall the outcome that followed a given response during the learning phase" (P.14). They found that the likelihood of selecting brand O3 (associated with positive images) over O1 (associated with negative images) was almost ten times lower if participants failed the contingency memory test than if they passed. Additionally, Gast et al. (2012) found that the US-CS conditioning was more successful in eliciting the CR if an individual could recall previously learnt pairings between US and CS. They separated the study into two sessions and examined participants' ability to recall US-CS pairings after several days in session two. The result suggested that the conditioning was more effective with those who could recall the pairings. The present study also added a memory test in the training session following every three pairings. The purpose of the present study was not to analyse the influences of a memory test in the SPOP and it was not designed to do so. Therefore, future studies might be required to evaluate this influence as it may be an effective addition to the procedure. If the memory test would be used as an assessment tool, increasing the difficulty of the test (e.g., involving more items) may help to avoid ceiling effects.

This study had several limitations. Firstly, the researcher selected the symbols and the neutral pictures for the experiment without any investigation into their neutrality. Therefore, for some participants, those symbols may have already had some behavioural function prior to the experiment. For example, Participants 18 pointed out that they ranked the drinks depending on if the symbol on the drinks looked safe. For example, he felt one of the symbols (¤) was unsafe because it looked like a skull symbol on poison. Although this symbol was associated with negative

pictures, it was unclear if the pairing or the participant's previous experience (or both) was responsible for their preference assessment outcomes and this statement. Furthermore, the same stimuli were used for all participants and, therefore, the function of the positive and negative images likely differed among participants. Most participants' body language or expressions (for example, covering their eyes when saw unpleasant images or making a sound such as "Eww!") suggested that some of them felt uncomfortable viewing the negative pictures. During the training session, however, Participant 20 indicated that he was interested in those negative images, such as the open-heart surgery picture. Hence, he may not have had the same negative feeling towards negative pictures as most other participants did. He achieved 97.92 % of accuracy on the memory tests, which was higher than most of the participants

Tyron and Cicero (1989) also pointed out that researchers may have a different perception about the stimuli from some of the participants. Therefore, for future research, it may be a good idea for the researcher to ask the participants to rank the symbols before the experiment begins and to compare the pre-ranking results with the post-ranking. Barnes-Holmes et al. (2000) used a pre- and post-rating procedure to resolve this issue. Alternatively, the researcher could do a short survey to see if participants have any thoughts related to these symbols or ask them if they have had experience with these symbols before. In the study of Hughes et al. (2016) as well as Staats and Staats (1957), they added a stimulus pre-ratings test. In the test, they asked the participants to evaluate a series of fictitious brand and images before the training and let the participants pre-rate the stimuli and asked them if they had previous experience with the stimuli. Many studies on stimulus equivalence or relational frame theory employ Chinese characters, but this may be problematic for the same reasons. For example, according to Clayton and Hayes (2004), they used 18 randomly selected Chinese characters as stimuli to examine the effectiveness of MTS and respondent-type (ReT) procedures for transfer of stimulus function. They recruited 13 students, and one of the participants was Japanese. The Japanese participant achieved 100% accuracy in both MTS and ReT test. As the Japanese participant stated, he/she was familiar with the stimuli, which may have facilitated their performance.

An additional shortcoming of this study is that, for some participants, after they read the information sheet, they would start to inquire about the theory of the experiment. Some other participants conjectured about the purpose of the study after they finished the experiment. What is more, some of them even discussed their understanding of the experiment when they were participating. Participants might

have inferred that the purpose of the research was to teach them to choose the symbols that were followed by "good" pictures. This phenomenon was called "good participant" by Hughes et al. (2016). As they described, a relevant detailed task instruction and research background introduction may imply to participants the purpose of the experiment. As a result, it is very likely that participants will make a presumption about the goal of the experiment and behave in a way that conforms to that goal (Hughes et al., 2016). In order to find out if response bias moderated the test result, Hughes et al. (2016) conducted "demand compliance" and "hypothesis awareness" checks. During the inspection, participants were asked about their conjectures on the intention of the experiment. The same limitation was mentioned by Barnes-Holmes et al. (2000). In the study, they suspected that a detailed introduction of the experiment may have caused the participants to guess the purpose of the tester and to cater to the tester's expectation. Therefore, for researchers who want to improve the test procedure or do further investigation in this field, adding a semi-structured interview may help researchers to determine if response bias has impacted on the outcome of the experiment. Alternatively, they may resort to some form of deception, provided that ethical guidelines are not violated.

Another limitation of the current research was that soda water was used in the experiment for participants to clean their mouths during the preference assessment. However, according to most of the participants, they felt unaccustomed to the taste of soda water. Additionally, the taste of soda water may have changed the taste of the soft drink, which would have changed the subjective taste of subsequent samples of cola. This may lead to bias in the preference test. However, the order of the brands was randomize across participants for the taste-test ranking assessment, which should have minimized this potential effect in the overall results.

Another limitation of this study is that a familiar drink, cola, was presented in the ranking assessment, which might have enabled participants to more easily identify that all of the “brands” were the same cola. Participant 17 reported that he could hardly tell the difference between the 6 cups of soft drink. Because all cups contained the same drink, which is a common tactic for a variety of research procedures, the participants might have guessed that they were the same, particularly after tasting them, and ranked them at random. If this occurred, then the taste-test ranking results may not be meaningful. In future research, providing participants with an unusual beverage, which they may not be familiar with, may help to avoid this potential issue.

This experiment required the participant to continuously attend to the

computer screen and, therefore, distractions and fatigue could be the main reasons for the low accuracy of some participants. For example, according to Participant 6 (one participant), during the test, she lost focus on the training because she parked her car in a tow away area and she was worried about her car being towed away. Therefore, she interrupted the experiment to move her car to a proper parking area. For Participant 5, in the first session, she checked her phone from time to time. After reminding her to stop watching the cell phone, her phone continued making noise. Participant 3 was easily distracted by background noise. According to her self-report, before she came for the experiment, she had just finished her work and felt tired. Participant 10 showed a little impatience in the second training for the same reason. Participant 18 had low attention levels as well. For example, he was impatient to listen to the instructions and began to do the training session without the researcher finishing the statement of the experimental description. During the experiment, he frequently looked away from the computer screen and turned back to talk with the researcher and discuss the experiment. This may be the cause of his low accuracy (33.33%) in the later MTS test. Therefore, for future studies, the researcher can draw on the experience of Hughes et al. (2016). In their study, they highlighted the importance of accuracy in the experiment. For example, a reminder sentence and percentage correct so far could be repeatedly shown on the computer screen. In this way, it may increase participants' concentration on the training.

Although before the experiment, all participants were aware some negative pictures would be used in the research and gave consent, some participants' reaction toward those negative pictures did have a negative influence on their performance. From the observation, Participants number 4, 6, 9 and 18 had a strong reaction to the negative pictures. Especially for Participant number 9, who sometimes diverted her eyes when those pictures were shown. This seemed to cause her performance to suffer in the first part of the experiment.

Despite the limitations, the study does have some important implications. One encouraging finding of this experiment is that unpleasant stimuli, such as those used in anti-smoking campaigns, appear to be more effective at driving choice than do pleasant stimuli, such as those used in common advertising of soft drinks and other harmful products. Therefore, the current research may contribute to the development of methods to encourage and help people to change certain bad habits such as cigarette smoking, binge drinking, drunk driving, net-addiction, and gambling. Some anti-smoking campaigns have used tactics similar to those evaluated in the current experiment. For example, governments in several countries now require that negative pictures of smoking-related disease be printed on the packaging of

cigarettes so as to let make smokers aware of the terrible consequences of smoking with the hope that this reduces their likelihood of purchasing or using these products. Some other anti-smoking campaigns used posters or damaged organs caused by diseases such as lung cancer and emphysema, wrinkles and stretch marks, and diabetic retinopathy to show people the damaging effect of smoking on health. In Queensland, Australia, an anti-smoking campaign used makeup technology to display the risk of long-term smoking. At the event, professional make-up artists did imitation makeup for young women in order to present the participants and everyone with the serious consequences of smoking – an ugly face. Participates were Women aged between 18 and 24. This activity aimed at emphasizing the negative impact of smoking on the human body. For example, smoking can gradually damage the smoker's hair, nails and skin. It causes the body to produce a substance, which accelerates the process of break downing of collagen and accelerates the ageing of the body. The chemicals in tobacco can damage the fibres in human skin; weaken the elasticity and toughness of the skin; accelerate the formation of wrinkles, and also enlarge the under-eye bags and cause sagging breasts (Murdaca, 2014). Similar strategies might also be applied to reduce binge drinking, drunk driving, smoking during pregnancy, net-addiction, drugs addict and gambling. Using cyber games as an example.

In China, many computer games begin with a statement such as, "addiction to games is bad for your health." However, the display time is relatively short, and the written words may have little impact on players. Therefore, it may be more effective to present aversive images related to computer game addiction directly to the players at the beginning of the game. Further research is required to understand the mechanisms that are responsible for the apparent success of some of these strategies, which would help to refine approaches to helping reduce harmful behaviour.

It seems that participants will select away from brands paired with negative images more than toward positive brands. Numerous studies had confirmed that overconsumption of sweet drinks was related to heart disease, obesity, diabetes, high blood pressure, dental caries and other diseases. For example, Fung et al. (2009) did a 24 years follow-up research (from 1980 to 2004) among women found that soft drinks were closely related to coronary heart disease. In addition, the research found that obesity caused by excessive sugar intake mediated the association between the consumption of soft drinks and the risk of coronary heart disease. Moreover, they found that artificially sweetened beverages had little impact on the risk of CHD. A worldwide study conducted by Singh et al. (2015) revealed that from 1980 to 2010,

by surveying more than 611,971 people across 51 countries, there were about a total of 184,000 deaths every year related to diabetes, heart disease and cancers, which caused by overconsumption of the sweet drinks. Among them, there were approximately 133,000 people die from diabetes, 45,000 people die due to heart disease and around 6,450 die from cancer. The researchers considered smoking, drinking, taking vitamins, family history, physical activity, body mass index (BMI) and other factors. The results showed that people who intake sugar-sweetened drinks every day would increase the chance of having a heart attack. Additionally, according to Sorensen et al. (2005), blood tests showed that frequent sugary drinking was associated with increased levels of C - reactive protein in the body, which induces inflammation, and inflammation was considered to be one of the leading causes of heart disease. According to Schulze et al. (2004), an eight-year survey of 50,000 nurses indicated that people who drinks one or more serving sweet drink in one day had an 80% greater chance for getting diabetes than non-drinkers. Most soft drinks contain large amounts of carbohydrates such as glucose, fructose and sucrose. The buccal bacteria can ferment these carbohydrates and produce acidic substances, which in return dissolve the enamel of the teeth, increasing the probability of tooth decay. The more people drink, the higher the risk of tooth decay (Marshall et al., 2003).

Printing negative pictures on soft drink packages is unlikely to gain public or government support, and it is unclear if it would be effective. An alternative strategy to printing graphical pictures directly on packaging may be to design a "logo" and then through public service announcements or mass media associate the logo with the diseases that are related to soft drinks. Then, the logo can be printed on relevant product packaging. In this way, when people encounter soft drinks and see the logo on the outer packing, the function of the negative images which have transferred to the logo may drive choice away from such products. Such a logo might function like a skull logo printed on toxic substances, but if associated with relevant negative images, it may meet with less resistance than a “scare tactic” such as placing a skull on the product. Similar “logos” could be developed for other harmful products and paired with other relevant pictures. The current research did not explore the effectiveness of such an approach relative to the approach of printing pictures directly on a product. Instead, a comparison was made between direct and indirect pairing in order to examine if the indirect pairing is more effective than direct pairing procedure to change preference with such a “logo.” Therefore, additional research is required to determine the relative effectiveness of this approach.

Although in the present research, the “direct pairing” approach appeared to be more effective than an “indirect pairing” approach, there may be an advantage to using an indirect pairing approach in some cases. According to Glock et al. (2012), warning labels and pictures of smoking-related diseases printed on cigarette packages may be ineffective in reducing smoking behaviour because individuals may perceive those negative pictures or warning statements as "external sources (information provided by other people)" (p.254). From a behavioural perspective, verbal statements to this effect might result in “counter control” by the individual whose behaviour could be modified by this intervention (Baum, 2005 p. 226). A direct pairing procedure may be more likely to provoke defensive responses, or “counter control.” Therefore, additional research is required to explore the effectiveness of a “logo” or similar tactics on preference for consumption of unhealthy products.

References

Anarchi.st (2017, May). *McDonald's Dead Dad Advert Commercial* [Video file]. Retrieved from https://www.youtube.com/watch?v=S1XM4INk8l8

Barnes-Holmes, D., Keane, J., Barnes-Holmes, Y., & Smeets, P. (2000). A derived transfer of emotive functions as a means of establishing differential preferences for soft drinks. *The Psychological Record, 50*(3), 493-511.

Barnes-Holmes, D., Barnes-Holmes, Y., Smeets, P., Cullinan, V., & Leader, G. (2004). Relational frame theory and stimulus equivalence: Conceptual and procedural issues. *International Journal of Psychology and Psychological Therapy, 4*(2), 181-214.

Baum, W. M. (2005). *Understanding behaviorism: Behavior, culture, and evolution (2nd ed.).* Malden, MA: Blackwell Pub.

Cicero, S., & Tryron, W. (1989). Classical conditioning of meaning—II. A replication and triplet associative extension. *Journal of Behavior Therapy and Experimental Psychiatry, 20*(3), 197-202.

Clayton, M., & Hayes, C. (2004). A comparison of match-to-sample and respondent-type training of equivalence classes. *The Psychological Record, 54*(4), 579-602.

Fung, T., Malik, V., Rexrode, K., Manson, J., Willett, W., & Hu, F. (2009). Sweetened beverage consumption and risk of coronary heart disease in women. *The American Journal of Clinical Nutrition, 89*(4), 1037-1042.

Gast, A., De Houwer, J., & De Schryver, M. (2012). Evaluative conditioning can be modulated by memory of the CS–US pairings at the time of testing. *Learning and Motivation, 43*(3), 116-126.

Hughes, S., De Houwer, J., & Perugini, M. (2016). Expanding the boundaries of

evaluative learning research: How intersecting regularities shape our likes and dislikes. *Journal of Experimental Psychology: General, 145*(6), 731-754.

Kinloch, J., McEwan, J., & Foster, T. (2013). Matching-to-sample and stimulus-pairing-observation procedures in stimulus equivalence: The effects of number of trials and stimulus arrangement. *The Psychological Record, 63*(1), 157-173.

Marshall T. A., Levy, S. M., Broffitt, B., Warren, J. J., Eichenberger-Gilmore, J. M., Burns, T. L., & Stumbo, P. J. (2003). Dental caries and beverage consumption in young children. *Pediatrics, 112*(3). Retrieved from http://pediatrics.aappublications.org

Mead, E., Cohen, J., Kennedy, C., Gallo, J., & Latkin, C. (2015). The role of theory-driven graphic warning labels in motivation to quit: A qualitative study on perceptions from low-income, urban smokers. *BMC Public Health, 15*, BMC Public Health, 2015, Vol.15.

Mothersbaugh, D. L. & Hawkins, D. (2015). *Consumer behavior: building marketing strategy.* United States: McGraw Hill Higher Education. ISBN: 1259232549.

Murdaca, C. (2014). Fans accuse Rachael Finch of 'airbrushing' pics from anti-smoking shoot. *TheFix.* Retrieved from https://thefix.nine.com.au

National Cancer Institute (2008). *The Role of the Media in Promoting and Reducing Tobacco Use. Tobacco Control Monograph No. 19.* Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute. NIH Pub.

Pornpitakpan, C. (2012). A critical review of classical conditioning effects on consumer behavior. *Australasian Marketing Journal (AMJ), 20*(4), 282-296.

Schulze, M. B., Manson, J. E., & Ludwig, D. S. (2004). Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women. *ACC Current Journal Review, 13*(11), 34-35.

Singh, G. M., Micha, R., Khatibzadeh, S., Lim, S., Ezzati, M., & Mozaffarian, D. (2015) Estimated global, regional, and national disease burdens related to sugar-sweetened beverage consumption in 2010. *Circulation, 132* (8), 639-666. https://doi.org/10.1161/CIRCULATIONAHA.114.010636

Sorensen, L., Raben, A., Stender, S., & Astrup, A. (2005). Effect of sucrose on inflammatory markers in overweight humans. *The American Journal of Clinical Nutrition, 82*(2), 421-427.

Smeets, P. M., & Barnes-Holmes, D. (2003). Children's emergent preferences for soft drinks: Stimulus-equivalence and transfer. *Journal of Economic Psychology, 24*(5), 603-618.

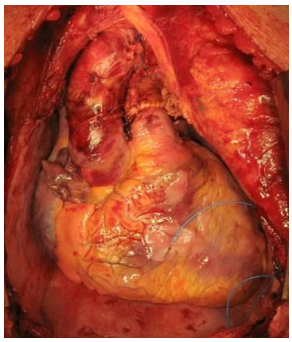
Staats, C. K., & Staats, A. W. (1957). Meaning established by classical conditioning. *Journal of Experimental Psychology, 54*(1), 74-80.

Tryron, W., & Cicero, S. (1989). Classical conditioning of meaning—I. A replication and higher-order extension. *Journal of Behavior Therapy and Experimental Psychiatry, 20*(2), 137-142.

Valdivia-Salas, S., Dougher, M., & Luciano, J. (2013). Derived relations and generalized alteration of preferences. *Learning & Behavior, 41*(2), 205-217.

Appendix 1: Pictures and Symbols

Negative pictures:

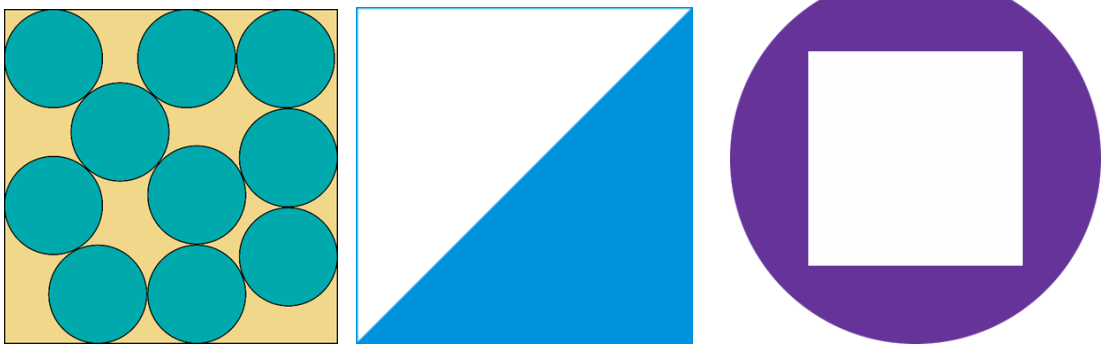


Positive pictures:



(Appendix 1 continued)

Neutral pictures:



Symbols:

S1

S2

S3

S4

S5

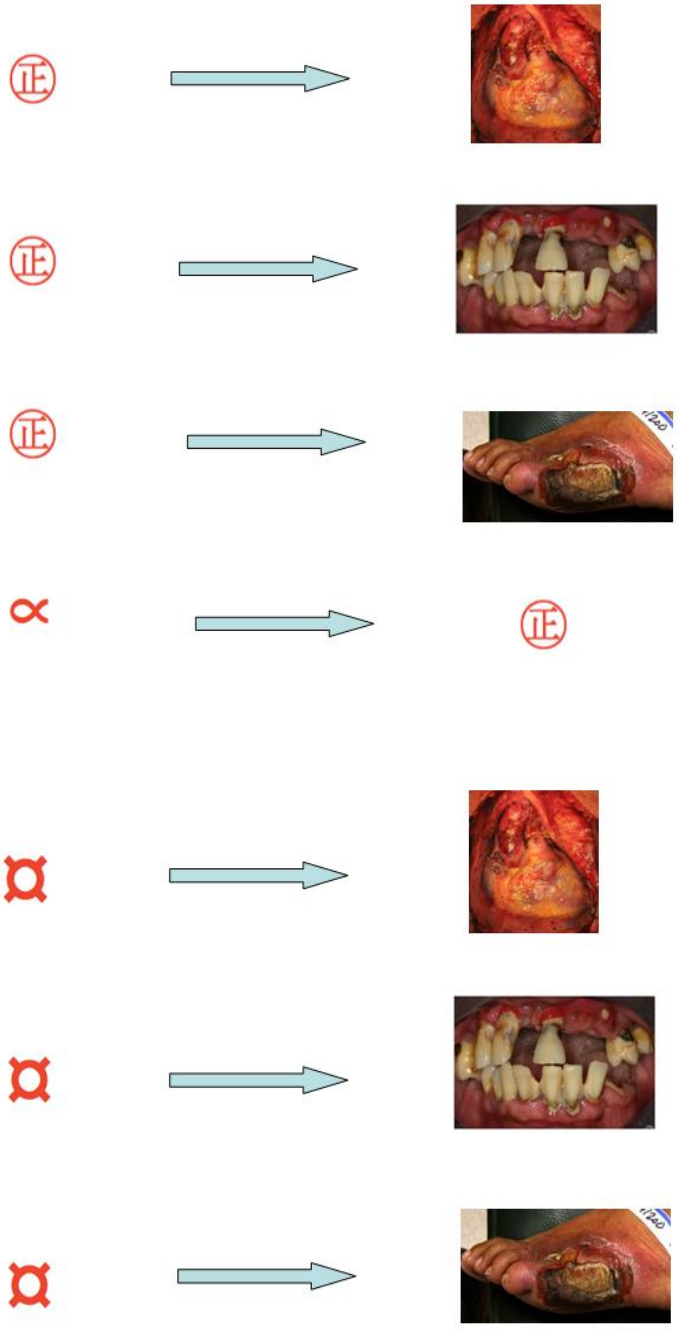
S6

S7

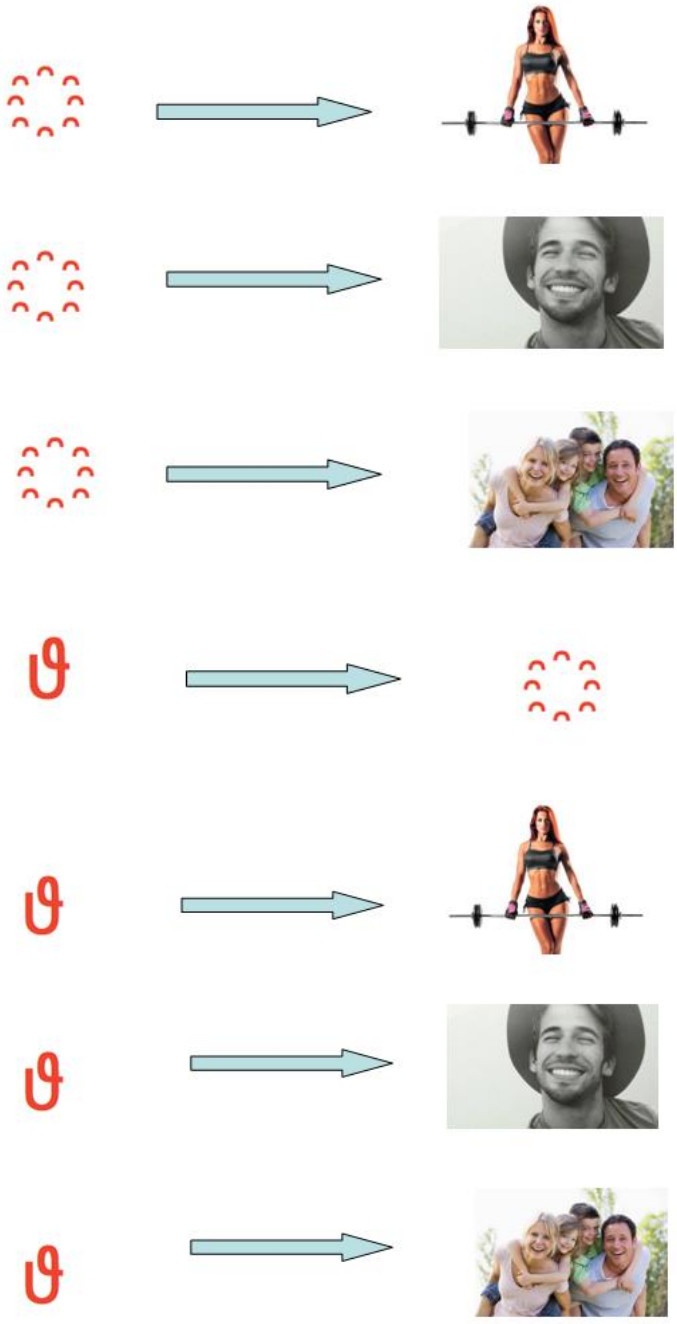
S8

S9

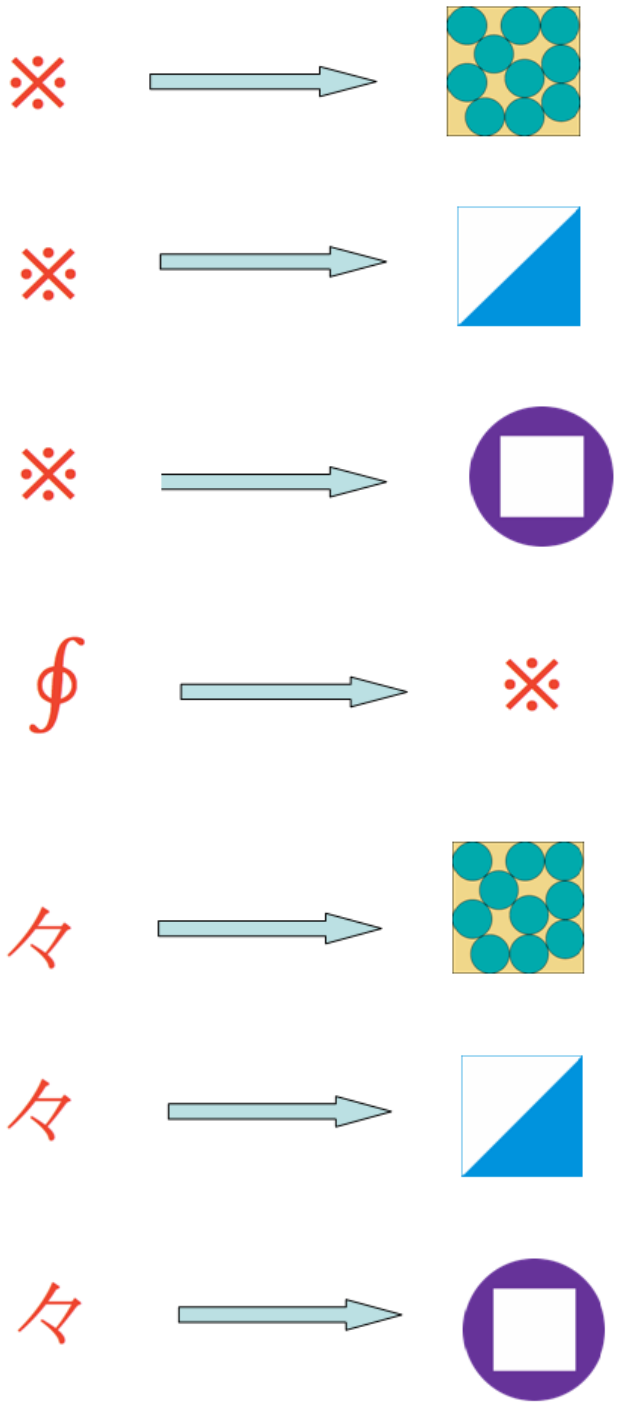
Appendix 2: Stimulus Pairs



(Appendix 2 continued)



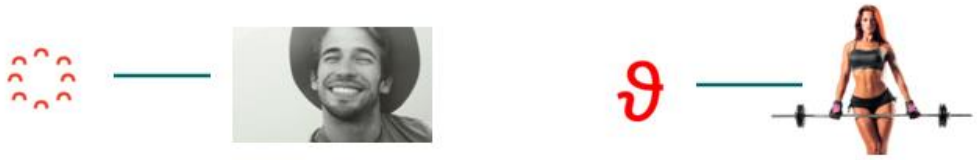
(Appendix 2 continued)



Appendix 3: Memory Test Example

Example of memory test

Instruction: Please choose one pair that you have not seen in the previous three pairings.



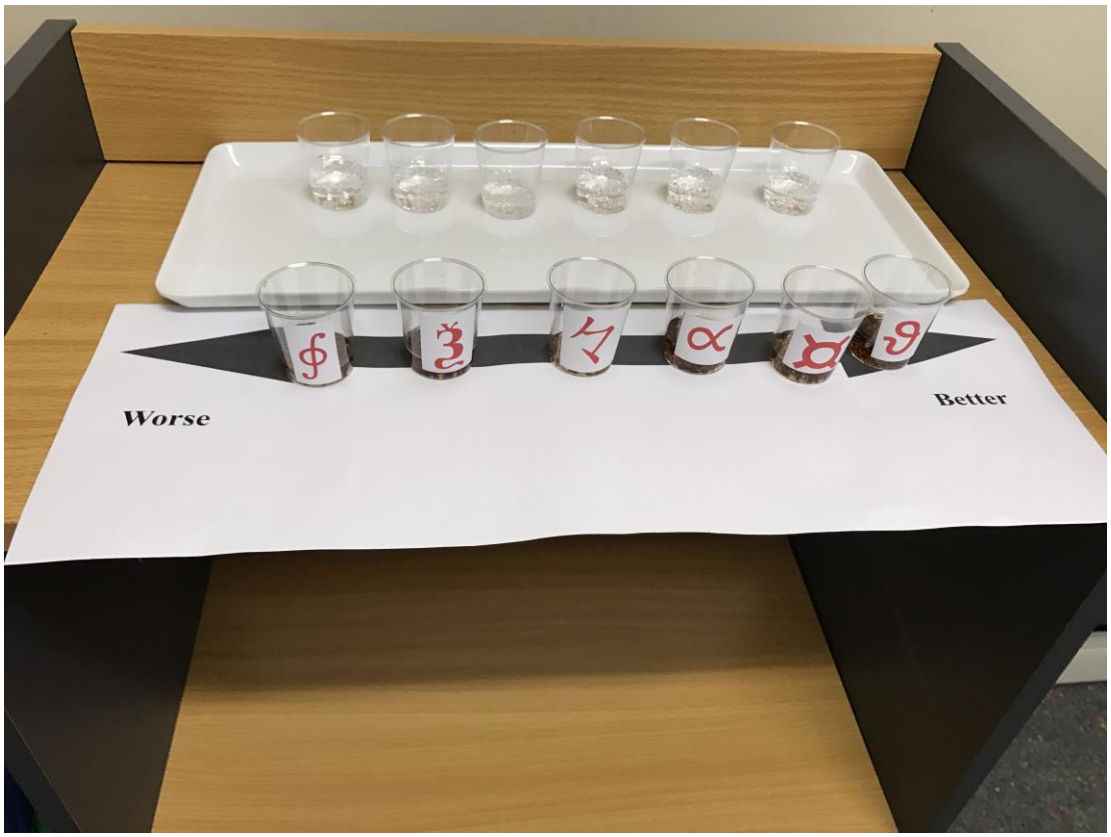
Appendix 4: Preference Assessment Data Sheet

(Example with data)

Please point to the one you prefer.

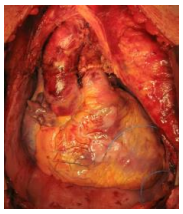
Participant's NO. _____01_____				**Participant's NO. ____02_____**		
	Trial 1	Trial 2	ranking	Trial 1	Trial 2	ranking
	3	4	4	5	4	1
	5	6	6	1	3	5
	4	3	3	2	1	6
	6	5	2	6	5	2
	1	1	1	4	6	4
	2	2	5	3	2	3

Appendix 5: Taste-test Ranking Scale



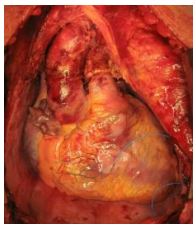
Appendix 6: Matching-to-Sample Test

Trial 1 - sample stimulus



The sample stimulus was centered in the top half of the screen and remained on the screen by itself for 1.5 s.

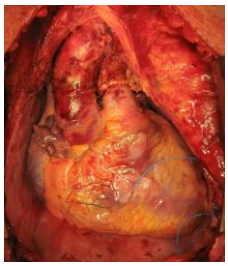
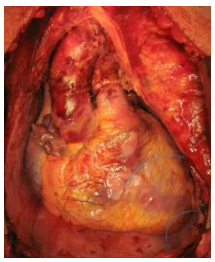
Comparison stimuli



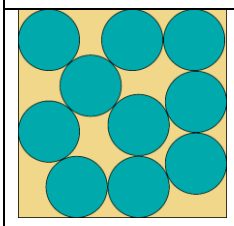
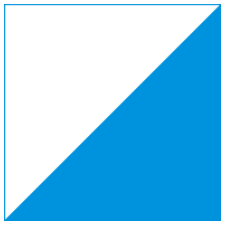
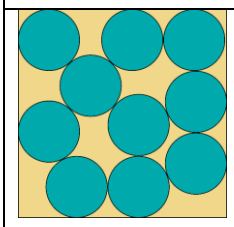
The participants are required to make their choice by single-clicking the correct stimulus.

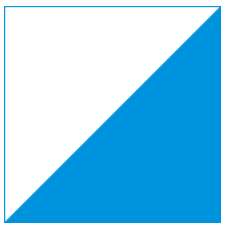
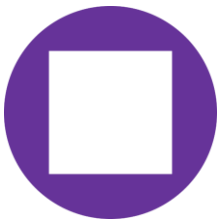
The sample and comparison stimuli remain on the screen until a response is made.

Appendix 7: Matching-to-Sample Trials

Sample stimulus	Comparison stimuli		
	Correct	**Incorrect 1**	**Incorrect 2**
	 Indirect negative		
BP1	S2	S5	S8
	 Indirect negative		
BP2	S2	S5	S8
	 Indirect negative		
BP3	S2	S5	S8
	 Direct negative		
BP1	S3	S6	S9
			

	Direct negative		
BP2	S3	S6	S9
	¤ Direct negative	ϑ	々
BP3	S3	S6	S9
	ǯ Positive indirect	∝	∮
GP1	S5	S2	S8
	ǯ Positive indirect	∝	∮
GP2	S5	S2	S8
	ǯ Positive indirect	∝	∮
GP3	S5	S2	S8
	ϑ Direct positive	¤	々

GP1	S6	S3	S9
			
GP2	S6	S3	S9
	 Direct positive		
GP3	S6	S3	S9
	 Indirect neutral		
NP1	S8	S2	S5
	 Indirect neutral		
NP2	S8	S2	S5
	 Indirect neutral		
NP3	S8	S2	S5
	 Direct neutral		
NP1	S9	S3	S6

	々 Direct neutral	¤	؏
NP2	S9	S3	S6
	々 Direct neutral	¤	؏
NP3	S9	S3	S6