INTERNATIONAL JOURNAL OF PSYCHOLOGY AND PSYCHOLOGICAL THERAPY 2011, 11, 3, pp. 377-388

# Exploring Screen Presentations in the Implicit Relational Assessment Procedure (IRAP)

**Claire Campbell*[1], Yvonne Barnes-Holmes[2], Dermot Barnes-Holmes[2] and Ian Stewart[3]**

[1]*Lancaster University, United Kingdom* [2]*National University of Ireland-Maynooth, Ireland*

[3]*National University of Ireland-Galway, Ireland*

## ABSTRACT

The current study attempted to systematically manipulate stimulus presentations in the Implicit Relational Assessment Procedure (IRAP) to determine the potential impact of this variable on implicit responding. The study comprised of four conditions that systematically manipulated the positions of the sample stimuli and the response options. Specifically, the Random-Random Condition randomized both sample stimuli and response options; Random-Fixed randomized sample stimuli, but response options remained in fixed locations; Fixed-Random Condition fixed sample stimuli but randomized response options; and Fixed-Fixed Condition fixed both sample stimuli and response options. The results demonstrated strong and predicted IRAP effects in all four conditions. Although the Random-Fixed presentation generated the strongest D-IRAP score, the randomization of the sample stimuli and response options were both critical to producing strong and significant D-IRAP scores because the Random-Random and Fixed-Random Conditions were only marginally smaller. The implications of the findings for existing and future research with the IRAP are discussed.

*Key words*: implicit attitudes, Implicit Relational Assessment Procedure (IRAP).

## RESUMEN

El presente estudio intenta manipular sistemáticamente presentaciones de estímulos en el *Implicit Relational Assessment Procedure* (IRAP) para determinar el impacto potencia de esta variable sobre el responder implícito. El estudio comprende cuatro condiciones que manipularon sistemáticamente las posiciones de la muestra de estímulos y las opciones de respuesta: condición Aleatorio-Aleatorio que aleatoriza tanto la muestra de estímulos como las opciones de respuesta; condición Aleatorio-Fijo que aleatorizó la muestra de estímulos pero con las opciones de respuesta en posiciones fijas; condición Fijo-Aleatorio que mantuvo fija la muestra de estímulos y aleatorizó las opciones de respuesta; y la condición Fijo-Fijo que mantuvo fijas la muestra de estímulos y las opciones de respuesta. Los resultados demuestran fuertes los efectos del IRAP predichos para las cuatro condiciones. Aunque la presentación Aleatorio-Fijo generó la mayor puntución D-IRAP, las aleatorizaciones de la muestra de estímulos y las opciones de respuesta fueron críticas para producir puntuaciones fuertes y significativas D-IRAP, ya que las condiciones Fijo-Fijo y Fijo-Aleatorio fueron sólo marginalmente menores. Se discuten las implicaciones de estos hallazgos para la investigación actual y futura sobre el IRAP.

*Palabras clave*: actitudes implícitas, Implicit Relational Assessment Procedure (IRAP).

* Correspondence concerning this article should be addressed to: Claire Campbell, Department of Psychology, Lancaster University, Bailrigg, Lancaster. UK LA1 4YF. E-mail: Claire.Campbell@lancaster.ac.uk

The Implicit Relational Assessment Procedure (IRAP) is an implicit measure that encompasses the principles of Relational Frame Theory (RFT), a modern behavioral approach to human language and cognition. The theory is built on the pivotal assumption that the behavioral units of language and higher-cognitive functioning are best defined in terms of *derived stimulus relations* (Hayes, Barnes-Holmes, & Roche, 2001). For RFT, derived relational responding is primarily *arbitrary* in nature and is governed by contextual cues via an appropriate history of multiple-exemplar training (see Barnes, 1996; Barnes, & Holmes 1991; Barnes, & Roche, 1996; Hayes, & Hayes, 1989). According to RFT, contextual cues facilitate various patterns of relational responding that are collectively referred to as *relational frames*. All relational frames possess the same three properties of mutual entailment, combinatorial entailment, and the transfer or transformation of stimulus functions, but each is identified according to the core type of stimulus relations involved. Some examples of relational frames include: coordination, opposition, distinction, hierarchy, and perspective-taking (Dymond, & Barnes, 1994).

The IRAP has demonstrated its utility in examining relational frames by including specific response options that are not included in other popular implicit methodologies. Consider a trial from the research by Barnes-Holmes, Barnes-Holmes, Power, Hayden, Milne, & Stewart (2006). On each trial, either PLEASANT or UNPLEASANT appeared as a sample stimulus at the top of the screen. The target stimulus in the middle of the screen comprised of a word from one of two concept categories that could be readily evaluated as positive (e.g., LOVE) or negative (e.g., CANCER). The relational response options were SIMILAR and OPPOSITE. In short, participants were required to indicate that the relationship between PLEASANT and LOVE is one of coordination (by choosing SIMILAR) or opposition (by choosing OPPOSITE).

Similar to the Implicit Associations Test (IAT), the IRAP predicts that already-established relations result in shorter response times than novel or less-established relations. For example, participants should take a shorter amount of time to relate the consistent relations PLEASANT/LOVE/SIMILAR than the inconsistent relations UNPLEASANT/LOVE/SIMILAR. This is known as an 'IRAP effect'. A range of studies have already demonstrated IRAP effects across an array of psychological phenomena, including attitudes to work and leisure (Chan, Barnes-Holmes, Barnes-Holmes, & Stewart, 2009), smoking (Vahey, Boles, & Barnes-Holmes, 2010), homosexuality (Cullen & Barnes-Holmes, 2008), self-esteem in children (Scanlon, 2008) and sexual offending (Dawson, Barnes-Holmes, Gresswell, Hart, & Gore, 2009). In addition to this there is also research that suggests that the IRAP is resistant to faking (McKenna, Barnes-Holmes, Barnes-Holmes, & Stewart, 2007).

Several studies in the cognitive literature have provided empirical support for the view that screen presentation plays an important role in implicit technologies. For example, Parris, Sharma, and Weekes (2007) demonstrated that coloring only a single letter rather than the full word in the Stroop Task may reduce or eliminate the Stroop interference effect. Furthermore, even elements of screen presentation that are *irrelevant* to the task may influence a participant's performance in subsequent experiments (Jiang & Leung, 2004; Pothos, 2005). For example, Deroost (2006) reported that the randomization

of visual or auditory stimulus sequences impaired reaction times to the randomized modality, as well as reaction times for modalities that were unaltered.

Implicit measures are primarily latency-based with stringent requirements on speed in order to ensure that responding is completely implicit. In the context of this paper we have adopted the definitions of Hughes, Barnes-Holmes and de Houwer (in press) by defining implicit responding as responding that is immediate, automatic, and non-declarative and by defining explicit responding as responses that are deliberate and controlled on the part of the participants. Latency-based measures are among those most commonly used to assess implicit responding because they provide an efficient and precise means to collect data about participants' indirect responding. These measures work on the principle that concepts that are consistent with a participant's beliefs will be linked in memory (are part of an existing relational network). It should, therefore, be possible to match these concepts more quickly than matching concepts that are inconsistent with beliefs (not part of an existing relational network). Because participants undertaking a latency-based implicit measure respond at sufficient speed to prevent them from manipulating their own responding, they are therefore heavily dependent on specific features of the stimulus presentation. This dependency may affect IRAP performances more than other implicit measures, because all IRAP blocks are identical in format. Indeed, all available IRAP studies have employed a presentation format in which both the sample stimuli and response options randomly switch positions across trials within each block. However, there is no empirical evidence to suggest the superiority of this format over any others (e.g., keeping either of these features fixed).

The current study attempted to systematically manipulate the presentation of sample stimuli and response options in the IRAP to determine what influence this might have on responding. Naturally, a simple IRAP (rather than one on race or nationality, for example) was chosen for this purpose. That is, participants completed a simple IRAP that required them to relate positive and negative target stimuli to the samples PLEASANT and UNPLEASANT using the response options SIMILAR and OPPOSITE. In line with previous research, participants also completed a questionnaire as an explicit measure of their attitudes toward the IRAP stimuli.

Participants were randomly assigned to one of four conditions, each of which systematically manipulated the locations of the sample stimuli and/or response options on-screen. Specifically, the Random-Random Condition randomized both sample stimuli and response options (i.e., the first term Random refers to the sample stimuli and the second term Random refers to the response options). The Random-Fixed Condition randomized sample stimuli (Random), but the response options always remained in fixed locations (Fixed). The Fixed-Random Condition fixed the sample stimuli (Fixed), but randomized the response options (Random). Finally, the Fixed-Fixed Condition fixed both sample stimuli (Fixed) and response options (Fixed). The primary aims of the research were to investigate whether the randomization of the sample stimuli and/or the response options influenced the IRAP effect and which of the four possible screen presentations potentially produced the strongest IRAP effect.

## METHOD

### *Participants*

Sixty undergraduate students of the National University of Ireland, Maynooth (NUIM) aged between 18 and 26 years old (*M*= 22 years and 11 months) participated in the current study. All were experimentally naive and none received remuneration for their participation. Participants were randomly allocated to one of four experimental conditions (15 per condition), across which the on-screen stimulus arrangements were manipulated. Twelve participants failed to reach criterion during practice blocks, and as a result their data were excluded from analysis, thus leaving 48 participants (12 per condition).

### *Setting*

The current study was conducted in a quiet room in the Department of Psychology at NUIM. All participants conducted the study on an individual basis, and only one was present in the laboratory at any one time. In all cases, the Experimenter remained present in the laboratory throughout and interacted directly with participants only during instructional, but not test phases.

### *Apparatus and Materials*

Participants were presented with twelve 13-point Likert scales, one for each of the target words that would subsequently be presented in the IRAP. Participants completed all IRAP trials on a DELL desktop computer with a Pentium 4 processor. The procedure was delivered via a program written in Visual Basic (Version 6.0) that controlled all aspects of stimulus presentation and the recording of all participant responses. Minor software modifications were necessary to distinguish the IRAP presentations that comprised each of the four conditions.

Each IRAP trial presented one of two sample word stimuli, PLEASANT or UNPLEASANT, one of 12 target stimuli (taken from Greenwald, McGhee, and Schwartz, 1998), and the two response options SIMILAR and OPPOSITE (see Table 1 for stimulus arrangements).

### *Experimental Conditions*

The four conditions in Experiment 1 varied only in terms of the randomization algorithm that controlled the presentation of the sample stimuli and the response options (the target stimuli were always presented in the middle of the screen):

*Random-Random Condition*. The Random-Random Condition contained two features that were alternated across trials (i.e., they were random). Specifically, the order of the presentation of the sample stimuli (PLEASANT and UNPLEASANT) at the top of the screen alternated in a quasi-random manner throughout each block of trials. In addition, the left-right positions of the response options (SIMILAR and OPPOSITE)

*Table 1.* The Stimulus Arrangements Employed in the IRAP.

| Sample Stimulus 1 | Sample Stimulus 2 |
|---|---|
| PLEASANT | UNPLEASEANT |
| Response Option Consistent with Sample 1 | Response Option Consistent with Sample 2 |
| SIMILAR | OPPOSITE |
| Targets Stimuli Consistent with Sample 1 | Targets Stimuli Consistent with Sample 2 |
| CARESS | ABUSE |
| FREEDOM | CRASH |
| HEALTH | FILTH |
| LOVE | MURDER |
| PEACE | SICKNESS |
| CHEER | ACCIDENT |

at the bottom of the screen also alternated in a quasi-random manner. Hence, the label Random-Random refers to the simultaneous randomization of both samples (Random) and response options (Random, respectively).

*Random-Fixed Condition*. In this condition, the presentation of the samples was randomized as above, but the locations of the response options remained fixed or identical across all trials (hence the term Random-Fixed). The actual presentation of each response option on either side of the screen was counterbalanced across participants (i.e., for half, SIMILAR was always on the left and OPPOSITE on the right, with the reverse presented to the remaining half).

*Fixed-Random Condition*. In this condition, the sample stimuli remained fixed, such that one appeared on the first 12 trials of each block and the second sample was then presented in the remaining 12 trials. The sequencing of the two samples was also counterbalanced across participants (i.e., half saw PLEASANT first, then UNPLEASANT, and this was reversed for the rest). The response options alternated in a quasi-random manner within blocks as before.

*Fixed-Fixed Condition*. In this condition, the sample stimuli were static (i.e., their sequence was not random) and the response options were always in the same fixed locations (hence, the label Fixed-Fixed). Again, the sequencing of the sample stimuli and the response options were counterbalanced across participants.

### *Procedure*

Given its aim to explore screen presentation effects, the current experiment employed neutral stimuli. In addition, all participants completed 12 Likert scales indicating their feelings towards each of the 12 target stimuli before taking part in the experiment. This allowed the Experimenter to ensure the reliability of the experimental categorizations employed subsequently in the IRAP.

Participants rated each word from +6 (Extremely Pleasant) to -6 (Extremely Unpleasant). For example, participants were asked to "Please indicate the extent to which you find the word PEACE pleasant or unpleasant by circling the appropriate number." [Capital letters indicate stimuli that actually appeared during the IRAP.]

Prior to exposure to the IRAP, participants were given instructions explaining that the next part of the study would involve a computer-based task requiring them to make

speedy and accurate responses that may or may not correspond to their own beliefs. On each IRAP trial, four words appeared simultaneously on the screen. A sample stimulus, either PLEASANT or UNPLEASANT, appeared at the top, with a target stimulus presented in the center, and two response options, SIMILAR and OPPOSITE at the bottom left- and right-hand corners of the screen. An illustrative example of the four IRAP trial-types can be found in Figure 1. The trial-types represent the four combinations that result when the sample stimulus is either PLEASANT or UNPLEASANT and the target stimuli are either positive or negative. All of the stimuli remained visible until the participant pressed one of the response keys. The task involved choosing one of the two response options (SIMILAR or OPPOSITE). To choose the term on the left, participants pressed the "d" key, and to choose the term on the right participants pressed the "k" key. Look, for example, at the screen shot in the top right hand corner of Figure 1, pressing "d" indicated that PLEASANT and PEACE were SIMILAR and pressing "k" indicated that PLEASANT and PEACE were opposite.

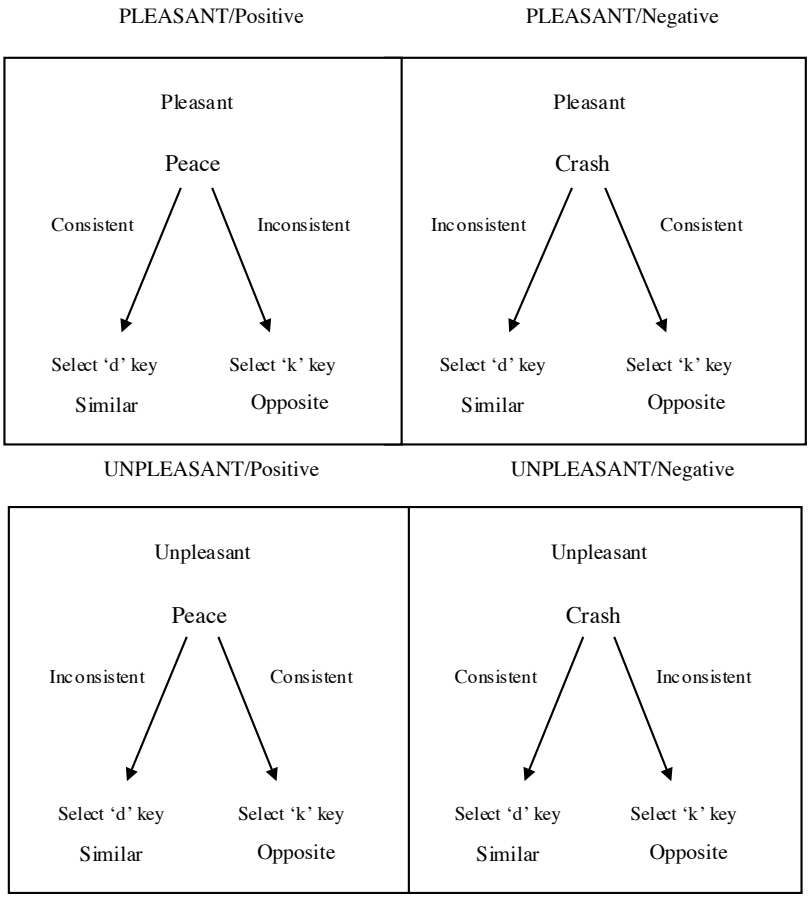


*Figure 1*. The four basic IRAP trial-types. The arrows, as well as the words 'Consistent' and 'Inconsistent' are not actually present on the screen during trials, but are included in the figure to indicate the type of trial in which that type of response would be considered correct.

At this stage of the research, minor variations in the instructions were provided for participants in each condition. Consider the following instructions employed in the Random-Random Condition:

> In this task the word at the top of the screen will randomly change between PLEASANT and UNPLEASANT. The words OPPOSITE and SIMILAR will also switch places randomly during the experiment. Sometimes the word OPPOSITE is on the right-hand side of the screen and SIMILAR is on the left-hand side of the screen and sometimes it is the reverse. So in this task you have to keep your eye on the word at the top (i.e., PLEASANT or UPLEASANT), the word in the middle (e.g., PEACE) and the places in which OPPOSITE and SIMILAR appear on the screen.

Although IRAP trials are generally referred to as test trials (as with the IAT), all incorrect responses are consequated with automated written *corrective feedback* and a correction procedure. That is, after an incorrect response was emitted the stimuli remained on-screen and a red X appeared directly below the target word. The X remained until the participant emitted a correct response, and the next trial then appeared automatically. Following a correct response, the stimuli were removed from the screen and the next trial was presented. No explicit feedback was presented for correct responding.

All blocks of *IRAP trials* (i.e., both practice and test blocks) were identical in format, and practice trials were always completed first. The number of practice blocks that participants received ranged from 2-8 and was contingent upon their performances therein. Once participants reached the mastery criteria (70% accuracy and 3000ms. response latency), they proceeded immediately to the first test block. All participants were exposed to a total of six test blocks. At the end of each block of trials, the IRAP presented participants with automated feedback on the percentage of trials correct and the median response time (in ms.) achieved during that block.

Each block of IRAP trials was designated, for experimental purposes, as *consistent or inconsistent*. Participants were exposed to a minimum of two practice blocks (one consistent and one inconsistent) and six test blocks (three consistent and three inconsistent). The IRAP sequence was always presented as alternating blocks of consistent and inconsistent trials. As a result, participants were required to switch their patterns of correct responding across blocks (i.e., the contingencies were reversed). In order to control for potential order effects, the sequencing of the blocks was counterbalanced across participants. That is, half of the participants were presented with a consistent practice block first, followed by the inconsistent practice block, followed by a consistent test block, and so on. In contrast, the other half were presented with an inconsistent practice block first, and so on.

The recording of a response on any trial as correct or incorrect depended on whether the trial had been categorized as consistent or inconsistent. Trials were referred to as *consistent* when the relations among the sample and target stimuli were consistent with the views believed to be held by participants prior to the study. That is, on these trials participants should more readily relate positive evaluations with PLEASANT, and negative evaluations with UNPLEASANT. Consider the trial presented in Figure 1. Correct responses on consistent trials required participants to respond PLEASANT/

PEACE/SIMILAR. Trials were referred to as *inconsistent* when the relations among the sample and target stimuli were *not* consistent with views attributed to the participants. On inconsistent trials correct responses involved selecting PLEASANT/PEACE/OPPOSITE. This combination generated four basic IRAP trial-types (see Figure 1).

The end of the sixth test block marked the end of the experiment for all participants. At this point, all were debriefed and thanked for their participation.

## RESULTS

The primary datum in the IRAP was response latency, defined as the time in milliseconds (ms) that elapsed between the onset of the trial and a correct response. Although accuracy was also recorded on every trial, the accuracy data were simply employed as a screening mechanism to ensure that all data contained within the subsequent analyses comprised of scores greater than 70% accuracy in the practice and test blocks. All others were removed from the analyses. In this experiment, participants focused more on accuracy than speed and all data sets removed from the analysis reflected slow responding. Data from the practice blocks were not included in the analyses.

In line with previous analyses of IRAP data, the response latency data for each participant were transformed into D-IRAP scores (see Cullen, & Barnes-Holmes, 2008) using an adaptation of the D-algorithm by Greenwald, Nosek, and Banaji (2003). The steps involved in calculating the D-IRAP scores were as follows: (i) only response latency data from test blocks were used; (ii) latencies above 10,000 ms. were eliminated from the dataset; (iii) all data for participants were removed if they produced more than 10% of test block trials with latencies less than 300ms; (iv) an overall standard deviation for all trial was calculated; (v) two mean latencies were calculated, the overall mean for consistent and inconsistent trials; (vi) the difference score was calculated by subtracting the mean latency of the consistent trials from the mean latency of inconsistent trials; (vii) the difference scores was divided by the standard deviation calculated in step (iv), yielding one overall D-IRAP score.

Figure 2 presents the mean overall D-IRAP scores for each condition. All conditions generated D-IRAP scores that were in the predicted direction. The positive D-IRAP scores for each condition indicate that participants categorized the positive words as pleasant and the negative words as unpleasant. The Random-Fixed Condition showed the largest D-IRAP score, followed closely by Random-Random. The D-IRAP score for Fixed-Random was slightly smaller, and Fixed-Fixed was the lowest of all. These findings suggest that the largest D-IRAP scores were obtained with randomization of either the sample stimulus or the response options, and the effect becomes very small when both are fixed.

A mixed repeated measures 4x2 Analysis of Variance (ANOVA) was conducted with condition and sequence (consistent-first vs. inconsistent-first) as between-participant variables. The analysis revealed a marginally significant main effect for condition [$F$ (3, 40)= 2.839, $p$= .05, $\eta_p^2$= .175], but not for sequence ($p$= .439), and there was no significant interaction effect ($p$= .634). Post-hoc analyses (Fisher's) indicated that the

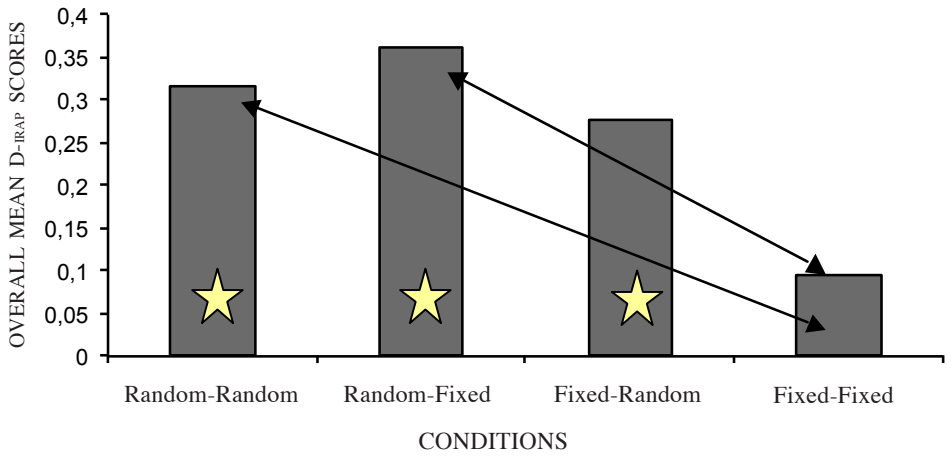


*Figure 2*. Mean overall D-IRAP score for each condition.Stars indicates that the D-IRAP score was significant relative to zero. Arrow indicates conditions that were significantly different from each other.

significant differences among conditions were between Random-Random and Fixed-Fixed ($p$= .023), and between Random-Fixed and Fixed-Fixed ($p$= .011), remaining $p$'s >.11.

Four one-sample t-tests were conducted to identify if the overall D-IRAP score for each condition differed significantly from zero. For three conditions, this was the case: Random-Random [$t$ (11)= 4.101, $p$= .002]; Random-Fixed [$t$ (11)= 4.307, $p$= .001]; and Fixed-Random [$t$ (11)= 3.035, $p$= .011]. The D-IRAP score for Fixed-Fixed did not differ significantly from zero ($p$= .126).

In line with other published IRAP studies, the results were also analyzed by trial-type. Four one-way ANOVA's were employed to identify if each trial-type differed across conditions. This analysis indicated that condition did not have a significant effect on responding across the PLEASANT/Positive, PLEASANT/Negative or UNPLEASANT/Negative trial-types ($p$ >.140). In contrast, condition did affect responding on the UNPLEASANT/Positive trial-type ($p$= .10). Post-hoc analyses revealed that this outcome occurred in the Fixed-Fixed Condition ($p$= .007, all other $p$'s >.617).

To assess the internal consistency of the IRAP, a split-half reliability score was calculated. This was based on two D-IRAP scores, one for odd trials and one for even trials. This was obtained in the same way as the overall D-IRAP score, except that the algorithm described previously was applied separately to all odd trials and to all even trials. The split-half correlation between odd and even scores, applying Spearman-Brown corrections, proved to be strong and significant, $r$= .644, $n$= 47, $p$ <.001. This indicated a reasonably strong level of internal consistency for the IRAP.

The Likert scales were designed to assess whether participants explicitly agreed with the experimental categorizations of the IRAP stimuli as positive or negative. The mean Likert ratings for the positive words ranged -4 to 6 with an overall mean of 4.8 (*SD* of 6.63). The mean Likert ratings for the negative words ranged from -6 to +6 with an overall mean of -4.48 (*SD* of 1.75). Hence, participants' evaluations of the two word groups were generally consistent with experimental categorizations.

## DISCUSSION

The basic aim of the current study was to determine the type of screen presentation that would yield the strongest IRAP effect. The key manipulation involved the presence or absence of the randomization of the sample stimuli and/or the response options. The results indicated that the randomization of the sample stimuli (with or without the randomization of the response options) generated the strongest D-IRAP scores. In addition, the D-IRAP scores generated by the randomization of the response options was not significantly different to those resulting from the randomization of the sample stimuli. This suggests that a certain degree of randomization is necessary to produce a significant and strong IRAP effect. The D-IRAP score that resulted from the fixed presentation of both samples and response options was negligible.

Practically all available IRAP studies have employed the Random-Random presentation format, in spite of the lack of empirical evidence to suggest the superiority of this format over the others. Although the current work indicated that the strongest IRAP effects were observed with the randomization of the sample stimuli but not response options (the Random-Fixed Condition), the difference between this and the Random-Random and Fixed-Random conditions was small and non-significant. This suggests that the consistent use of the Random-Random format in IRAP research to date has been wise, and has at least to some extent facilitated the strong IRAP effects reported therein. In line with these findings and the existing literature's use of this format, we would suggest that future IRAP studies continue to employ the Random-Random presentation, although the use of the Random-Fixed and Fixed-Random formats also offer viable alternatives. In any case, the current results also support those reported previously in the observation of strong and predicted IRAP effects.

Stimulus presentation plays an important role in any experimental methodology. For example, Bowe, Miller, and Green (1987) demonstrated that the quality and location of stimuli affects learning. Indeed, stimulus presentation is particularly important in implicit methodologies. This sensitivity results, at least in part, from the delicate balance between the need to make the task challenging (rather than turning it into an explicit measure because the time constraint is loose), while ensuring that the task is do-able. In short, high speed and accuracy criteria restrict the time a participant has to search around the screen and formulate a correct response. Indeed, several researchers have documented the importance of screen presentation in implicit tests (Deroost, 2006; Jiang & Leung, 2004; Pothos, 2005).

While no other IRAP study has manipulated screen presentations, several have examined other methodological features that potentially influence the size of IRAP effects. Most notable among these are several recent studies that have specifically manipulated the response time requirement in terms of reducing this from the original 3000ms to only 2000ms. For example, Barnes-Holmes, Murphy, Barnes-Holmes, & Stewart (2010) have reported stronger IRAP race effects (i.e., greater pro-White and anti-Black stereotyping) with the faster speed requirement. Taken together, this work highlights the utility of further research to determine which, and to what extent, specific IRAP features influence IRAP effects.

A potential limitation of the current study lies in the fact that the maximum response latency employed was 3000ms. and the accuracy criterion was 70%. This was the accepted convention at the time that this research was conducted. However, as mentioned above, in light of the work of Barnes-Holmes *et al*. (2010) it would seem important to replicate this study with the more stringent time and accuracy constraints. However, while the work of Barnes-Holmes *et al*. (2010) indicates that participants may still have some control over their responding at 3000ms., there is no evidence to suggest that one of the screen presentation formats investigated here would become superior over the others with increased speed and accuracy criteria.

The fixed or random nature of the stimuli presented in the IRAP also raises the related issue about the degree of flexibility that participants need to have, particularly in switching across blocks between consistent and inconsistent responding. The data here suggest that some randomization of sample stimuli or response options is necessary to produce strong and significant IRAP effects and that without randomization there is only a weak non-significant effect (e.g., Fixed-Fixed). Hence, one might assume that greater flexibility within the screen presentations *across trials* permits greater flexibility in responding *across blocks*, hence improved accuracy and/or speed.

There is recent empirical evidence that supports the importance of flexibility in this regard. For example, O'Toole and Barnes-Holmes (2009) used the IRAP to target similar/different and before/after relations in conjunction with the explicit Kaufman Brief Intelligence Test. Consistent with their predictions, participants who produced higher scores on the intelligence test responded more quickly on the IRAP. Specifically, the inconsistent IRAP trials produced a larger number of significant correlations with the explicit measure than consistent trials. In other words, participants who performed better on the intelligence test were not only faster at IRAP responding, but also demonstrated a greater degree of relational flexibility. It is reasonable to assume, therefore, that this flexibility across blocks may be facilitated by requiring considerable flexibility in responding to rapidly changing features of screen presentations across trials. The current findings support this view.

In summary, the current study demonstrated that of the four possible screen presentations employed all but one (the Fixed-Fixed format) went some way to facilitating the IRAP effect. The strongest IRAP effect was produced in the Random-Fixed Condition followed by the Random-Random Condition and the Fixed-Random condition. As there was no significant difference between the IRAP effects produced by each condition it would appear that some randomization of the sample stimuli or response options is critical to producing a strong and significant IRAP effect. These outcomes suggest that the consistent use of the Random-Random format in the IRAP research to date has been wise and has to some extent facilitated the strong IRAP effect commonly reported.

## REFERENCES

Barnes-Holmes D, Barnes-Holmes Y, Power P, Hayden E, Milne R, & Stewart I (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist, 32,* 169-177.

Barnes-Holmes D, Murphy A, Barnes-Holmes Y, & Stewart I (2010). The Implicit Relational Assessment Procedure (IRAP): Exploring the impact of private versus public contexts and the response latency criterion on pro-White and anti-Black stereotyping among white Irish individuals. *The Psychological Record, 60,* 57-80.

Barnes-Holmes D, Murtagh L, Barnes-Holmes Y, & Stewart I (2010). Using the Implicit Association Test and the Implicit Relational Assessment Procedures to measure attitudes towards meat and vegetables in vegetarians and meat-eaters. *The Psychological Record, 60,* 287-306.

Bowe CA, Miller JD, & Green L (1987). Qualities and locations of stimuli and responses affecting discrimination learning of chincillas (Chinchilla laniger) and pigeons (Columbia livia). *Journal of Comparative Psychology, 101,* 132-138.

Chan G, Barnes-Holmes D, Barnes-Holmes Y, & Stewart I (2009). Implicit Attitudes to Work and Leisure Among North American and Irish Individuals: A Preliminary Study. *International Journal of Psychology and Psychological Therapy, 10*, 453-474.

Cullen C & Barnes-Holmes D (2008). Implicit pride and prejudice: A heterosexual phenomenon? In TG Morrison and MA Morrison (Eds.), *Modern prejudice* (pp. 195-223). New York: Nova Science.

Dawson DL, Barnes-Holmes D, Gresswell DM, Hart AJP, & Gore NJ (2009). Assessing the implicit beliefs of sexual offenders using the Implicit Relational Assessment Procedure: A First Study. *Sexual Abuse: A Journal of Research and Treatment, 21,* 57-75.

Deroost N & Saetens E (2006). Spatial processing and perceptual sequence learning in SRT tasks. *Experimental Psychology*, *53*, 16-30.

Greenwald AG, McGhee DE, & Schwarz JLK (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464-1480.

Greenwald AG, Nosek BA, & Banaji, MR (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85,* 197-216.

Jiang Y & Leung A (2004). Implicit learning of ignored visual context. *Journal of Vision*, *4*,188.

McKenna IM, Barnes-Holmes D, Barnes-Holmes Y, & Stewart I (2007). Testing the Fake-ability of the Implicit Relational Assessment Procedure (IRAP): The First Study. *International Journal of Psychology and Psychological Therapy, 7*, 253-268.

O'Toole C & Barnes-Holmes D (2009). Three chronometric indices of relational responding as predictors of performance on a brief intelligence test: The importance of relational flexibility. *The Psychological Record, 59,* 119-132.

Parris BA, Sharna D, & Weekes D (2007). An optimal viewing position effect in the Stroop task when only one letter is in the color carrier. *Experimental Psychology, 54,* 273-280.

Pothos EM (2005). Expectations about stimulus structure in implicit learning. *Memory and Cognition, 33*, 171-181.

Scanlon G (2008). *The utility of the Implicit Relational Assessment Procedure (IRAP): Measuring self-esteem profiles of children with Special Educational Needs and assessing the attitudes of teachers to children with Emotional Behavioural Difficulties in mainstream education*. Unpublished doctoral thesis, Department of Psychology, National University of Ireland Maynooth.

Schmidtke V & Heuer H (1997). Task integration as a factor in secondary-task effects on sequence learning. *Psychological Research, 60,* 53-71.

Vahey N, Boles S, & Barnes-Holmes D (2010). Measuring Adolescents' Smoking-related Social Identity Preferences with the Implicit Relational Assessment Procedure (IRAP) for the First Time: A Starting Point that Explains Later IRAP Evolutions. *International Journal of Psychology and Psychological Therapy, 9*, 317-334.

*Received, March 5, 2010*
*Final Acceptance, June 5, 2011*