Effects of Perspective Taking and Values Consistency in Reducing Implicit Racial Bias*

Efecto de la Toma de Perspectiva y Consistencia en Valores en la Reducción del Sesgo Racial Implícito

Carmen Beck[a]
The Chicago School of Professional Psychology, United States
ORCID: https://orcid.org/0000-0002-8569-3090
Yors Garcia
Pontificia Universidad Javeriana, Colombia
ORCID: https://orcid.org/0000-0002-8868-0682
Robyn Catagnus
The Chicago School of Professional Psychology, United States
ORCID: https://orcid.org/0000-0002-4135-568X

[a]Correspondence author. Email: cbeckresearch@gmail.com

ABSTRACT

This study explored the effect of perspective taking and values consistency tasks on reducing implicit racial bias. Using a repeated measures design with a control group, 39 participants, 20 female and 19 males, 18-54 years old, who identified as White were administered the Implicit Relational Assessment Procedure. All participants in the experimental group ($n = 19$) and control group ($n = 20$) completed the Toronto Empathy Questionnaire, Modified Modern Racism Scale, Valuing Questionnaire, and a Likert scale. Experimental group participants completed brief values consistency and perspective taking tasks whereas the control group completed a guided task. A 2 x 4 mixed repeated measures analysis of variance ANOVA was conducted to determine whether there was any interaction effect between the group and trial types and a MANOVA to identify differences in the explicit measures between both groups. Results show that after the values work and perspective taking exercises, participants in the experimental group recorded shorter mean responses for Inconsistent-Black trial blocks versus Consistent-Black trial blocks as compared to the control group. Additionally, a statistically significant impact for interaction between condition and trial type was found for the Consistent-White trial type in the experimental group. Recommendations for future research are provided.

Keywords
racism; values; perspective taking; implicit relational assessment procedure; behavior analysis.

RESUMEN

Este estudio exploró el efecto de la toma de perspectiva y consistencia en valores en la reducción del sesgo racial implícito. Se utilizó un diseño de medidas repetidas con grupo control. A un total de 39 participantes, 20 mujeres y 19 hombres entre 18 y 54 años, se les administró el Procedimiento de Evaluación Relacional Implícita. Todos los participantes del grupo experimental ($n = 19$) y del grupo control

(*n* = 20) completaron el Cuestionario de Empatía de Toronto, Escala de Racismo Moderno Modificada, Cuestionario de Valores y una escala Likert. El grupo experimental realizó varios ejercicios en valores y toma de perspectiva, mientras que al grupo control se le asignó una tarea guiada. Se realizó un ANOVA mixto de medidas repetidas (2 x 4) para determinar el efecto de interacción entre grupo y tipos de prueba, y un MANOVA para identificar diferencias en las medidas explícitas entre ambos grupos. Los resultados mostraron que el grupo experimental registró respuestas más cortas para los bloques de prueba Inconsistente-Negro en comparación con el grupo control. Además, se encontró una diferencia estadísticamente significativa para la interacción entre condición y tipo de prueba en el bloque de prueba Consistente-Blanco del grupo experimental. Se presentan recomendaciones para futuras investigaciones.

Racial inequalities all over the world have been analyzed from the perspective of different psychological approaches for decades without a final resolution (Roberts & Rizzo, 2021; Salter et al., 2018). An explanation could be the conceptual and theoretical differences in describing stereotyping, prejudice, discrimination, and racism. For example, current psychological research has embraced the modern anthropological perspective that race is a social, rather than a biological, category (Yudell et al., 2016). In other words, the definition of race is influenced by factors of cultural importance such as focusing on visual cues like skin color and facial features to categorize people into groups (Obasogie, 2013). Conversely, the field of behavior analysis, a natural science that seeks to understand the behavior of individuals, has been concerned with analyzing the environmental and verbal contingencies supporting existing racial inequalities (Matsuda et al., 2020; Mizael & Sampaio, 2019; Saini & Vance, 2020). Said differently, whereas the field of psychology has historically emphasized the *form* of racism within a culture, behavior analysis has focused on the *function* of racism by analyzing the environmental contingencies that maintain it.

The traditional behavior-analytic view states that racism is acquired through three forms of learning processes: respondent, operant, and observation (Matsuda et al., 2020). All these three learning processes describe behavior in terms of direct contingencies processes, that is, all behavior is shaped and maintained by the direct contact with environmental consequences and extended through topographical generalization. More recently, however, derived relational responding (DRR) (Hayes et al., 2001), has emerged as an additional behavioral process that goes beyond the direct contingency control. DRR is an empirical effect in which learning relations between arbitrary stimuli result in the emergence of several additional untrained relations. Take for example, when a child learns that living in a White community is safer than in a Latino neighborhood. Then without a direct exposure or contact with this group of individuals, she may say that Latino neighborhoods are dangerous. Two behavior-analytic approaches, stimulus equivalence and relational frame theory (RFT) (Albright et al., 2021; Fryling et al., 2020; Matsuda et al., 2020), have used DRR processes to describe the establishment, development and maintenance of racial views. Stimulus equivalence states that equivalence is a basic behavioral process that emerges automatically and immediately as a consequence of a certain training history, whereas RFT holds that relating is a generalized operant that is created by a history of reinforcement across multiple exemplars (Fryling et al., 2020; Hayes et al., 2001).

In the last 30 years, behavioral researchers have used stimulus equivalence and RFT to assess the acquisition and generalization of racial bias (de Carvalho & de Rose, 2014; Dixon et al., 2009; Watt et al., 1991), and attempted to modify it in the laboratory (Dixon & Lemke, 2007; Mizael et al., 2016). Overall, findings in these studies have shown that racial bias once established, becomes resistant to extinction. Furthemore, within the last 16 years, RFT researchers have evaluated an alternative approach to assess the development and maintenance of racial bias using a method derived from RFT research named the Implicit Relational Assessment Procedure

(IRAP) (Barnes-Holmes et al., 2020). The IRAP is a computer-based, reaction time task. Its procedural parameters have been discussed in detail in many other papers (Barnes-Holmes et al., 2020; Hussey et al., 2015), and so only a brief overview will be provided here. Participants completing this computer-based task were asked to confirm or deny particular sets of relations under time pressure in accordance with particular rules for responding. If they showed higher levels of fluency (i.e., a combination of both speed and accuracy) towards one relational pattern than another, then this suggested that the latter pattern was more strongly established and thus more influential in their repertoire. For example, a participant might be displayed the stimuli "Latino people" and "dangerous" and be required to respond "Same" during one block trial and "Opposite" in another block. If participants respond faster to "Same" than to "Opposite" on this particular trial type, then, this might be taken as a measure of response bias toward Latino people and used to predict responding in another context (Barnes-Holmes et al., 2020).

A small number of studies (Barnes-Holmes et al., 2010; Drake et al., 2015; Dunne et al., 2018; Power et al., 2017a, 2017b) have demonstrated the usefulness of the IRAP to identify a response bias. Although response bias is sometimes referred to as racial bias, they are distinct in that response bias refers to a behavioral pattern observed in IRAP studies, whereas racial bias is used as a "proxy for a mental construct or implicit attitude in a cognitive or social psychology sense" (Barnes-Holmes et al., 2020, p. 267). To illustrate, Barnes-Holmes et al. (2010) displayed participants one of two label stimuli ("Safe" and "Dangerous") on each trial with a picture of a white or black man holding a gun as a target stimulus. During the IRAP task participants were required to respond on some block trials pro-white and anti-black (e.g., pressing a key for "True" rather than "False" when "Safe" appeared with a picture of a white man). On other trial blocks they were required to respond pro-black and anti-white (e.g., pressing a key for "True" rather than "False" when "Safe" appeared with a picture of

a black man). Results showed that participants responded "True" more quickly than "False" when displayed "Dangerous" and pictures of black men holding a gun; however, when pictures were of white men holding guns, participants responded "False" more quickly than "True". In other words, a response bias, responding more favorably (faster) to white men holding guns seemed consistent with the participant's pre-experimental histories. Additional IRAP studies have shown similar findings consistent with this response bias pattern (Drake et al., 2015; Dunne et al., 2018; Power et al., 2017a, 2017b).

A growing number of IRAP studies have explored the usefulness of this tool to assess implicit bias in multiple social domains, for example, national identity, religion, race, gender, sexuality and sexual preferences, age, body image, and smoking as stigmatized behaviors (Barnes-Holmes et al., 2020; Vahey et al., 2015). Despite the effectiveness of IRAP in identifying implicit bias in different social dimensions, there is a limited number of studies attempting to change or modify implicit bias (Ferroni et al., 2019; Kishita et al., 2014). In fact, recent research suggests that implicit bias is malleable and can be influenced, at least in the short term, by a number of interventions (FitzGerald et al., 2019; Forscher et al., 2019). For example, researchers have used mindfulness-based approaches, perspective taking, stereotype replacement, thought suppression, or value consistency to reduce racial bias, prejudice and discrimination. In this study two particular interventions are highlighted, perspective taking (Arauz et al., 2017; Garcia et al., 2021; Graham et al., 2015; West et al., 2013) and values clarification (Pashak et al., 2018; Todd et al., 2011; Wang et al., 2014). In behavior analysis these two interventions have been explored within acceptance and commitment therapy (ACT).

ACT is a behavior-analytic intervention grounded in functional contextualism, RFT, and evolutionary principles (Hayes et al., 2011). In the behavior-analytic field ACT is termed acceptance and commitment training (ACTr), to avoid confusion with how ACT is used in

traditional clinical psychology settings (Dixon & Hayes, 2022; Tarbox et al., 2022). The goal of ACT is to foster psychological flexibility, which refers to the ability to engage in values-based actions while in the presence of aversive functions. Psychological flexibility is composed of six interdependent skills, they are commonly known as defusion, acceptance, present moment attention, self-as-context, values, and committed actions (Hayes et al., 2011; Tarbox et al., 2022). Prior studies have indicated that perspective taking predicts anti-racist/anti-sexist behaviors across multiple domains (Davis et al., 2021). In laboratory-based studies, researchers have also shown that teaching perspective taking skills, for example, taking the perspective of another individual, produced a more positive interracial evaluation, increased awareness of racial inequality, and more approach-oriented behaviors toward a Black confederate (Todd et al., 2011; Todd et al., 2012). Similarly, Graham et al. (2015) observed that living consistently with ones' values moderated the relationship between exposure to racial stressors and depression and anxiety. Lastly, Lillis and Hayes (2007) reported that an ACT-based intervention led to significant changes in students' behavioral intentions, including their interest in seeking contact with students of other races or ethnicities, and joining diversity-related organizations.

Results from the above mentioned studies suggest the use of interventions that foster values and perspective taking, that is, taking specific actions toward marginalized groups as something, maybe, more effective than altering the content of prejudice or racial bias. Said differently, modifying behavioral patterns that sustain racial bias is maybe more effective than changing the content of thoughts that are indicators of racial bias. Although IRAP studies have successfully modeled implicit bias in the laboratory, there is a limited number of studies assessing the impact of ACT-based interventions such as values and perspective taking to reduce response bias using IRAP preparations. The goal of this study, therefore is, to extend prior research using ACT components to reduce racial bias toward Black people (Banks et al., 2021; Lillis & Hayes,

2007; Todd et al., 2012; Williams et al., 2020). Specifically, the IRAP will be used to determine whether implicit levels of bias have been reduced after the introduction of perspective and values clarification training.

Methods

Participants

Forty-one individuals (21 female, 20 male) aged from 18 to 54 years (M = 34.9, SD = 9.48) were recruited from two colleges and in the social media. The study was approved by the Institutional Review Board (IRB) of The Chicago School of Professional Psychology and Pepperdine University. Both institutions approved the study and sent research flyers through massive emails. In addition, the same research flier was posted in the main researcher's social media. All participants who met the inclusion criteria signed an informed consent. The informed consent outlined the expected duration of participation, the number of sessions required and how long the experiment should take. The informed consent also included information about how data will be stored and how long. Procedures to withdraw consent were also described. Many prospective participants were rejected due to the fact that the IRAP software is written in JAVA and is not compatible with devices that run iOS or Chrome OS. All participants were white U.S. citizens, they were randomly divided into two groups, experimental (21 participants) or control (20 participants). Two participants from the experimental group failed to reach the practice criteria (i.e., ≥ 80% correct and median response latency of ≤ 2000 ms) and did not continue to the test blocks phase. A total of thirty-nine participants completed all phases of the study. Participants spoke fluent English and had normal or corrected to normal vision and completed the study on a voluntary basis. Participants were not informed about the final goal of this project, namely, decreasing racial bias in a White sample.

Materials and Apparatus

Experimental sessions were conducted in an individual virtual setting (Zoom) with the first author. Participants used their own personal computer to access meetings and materials. Participants received hyperlink invitations for each meeting via email. The presentation of any session specific materials was shared via the One Drive folder in their email.

Implicit Relational Assessment Procedure (IRAP)

The IRAP is a computer-based program that records response latency to various stimulus relations. In this study, we used the GO-IRAP software. this is a freely available software (at https://go-rft.com/go-irap/). The label stimulus, either "White People" or "Black People", was displayed at the top of each trial along with a target stimulus, "reliable", "unreliable", "careful", "careless", "honest", or "dishonest". Participants were asked to respond, "True" or "False" using the "d" or "k" keys (Table 1). Correct responding was contingent upon pre-determined rules that were not known to the participants before the start of each trial block. This rule was reversed between each trial block. Half of the trial blocks required pro-White consistent responding (i.e., White-People-Reliable, White-People-Careful, White-People-Honest) while the second half required responses consistent with pro-Black bias (i.e., Black-People-Reliable, Black-People-Careful, Black-People-Honest). The order of trial blocks was counter-balanced across participants. Each of the 6 target words was displayed once with each of the two stimulus labels to produce 12 trials in each block. After each block, participants were displayed their median latency and mean accuracy (i.e., percentage of correct responses).

Table 1
Stimulus Arrangements Displayed by the IRAP

Label 1 White People	Label 2 Black People
Target 1 Target 1 #1: Reliable Target 1 #2: Careful Target 1 #3: Honest	Target 2 Target 2 #1: Unreliable Target 2 #2: Careless Target 2 #3: Dishonest
Response 1	Response 2
True	False
Rule A: Consistent	Rule A: Consistent
Label 1 White People	Label 2 Black People
Target 1 #1: "Reliable" + "True" Target 1 #2: "Careful" + "True" Target 1 #3: "Honest" + "True" Target 2 #1: "Unreliable" + "False" Target 2 #2: "Careless" + "False" Target 2 #3: "Dishonest" + "False"	Target 1 #1: "Reliable" + "False" Target 1 #2: "Careful" + "False" Target 1 #3: "Honest" + "False" Target 2 #1: "Unreliable" + "True" Target 2 #2: "Careless" + "True" Target 2 #3: "Dishonest" + "True"

Dependent Variable and Measurement

The Toronto Empathy Questionnaire (TEQ) (Spreng et al., 2009)

The TEQ was administered to participants via Google forms. The TEQ measures empathy and emotional processes. Item responses are scored according to Never = 0; Rarely = 1; Sometimes = 2; Often = 3; Always = 4. Positively worded questions include 1, 3, 5, 6, 8, 9, 13, and 16. Negatively worded questions were reverse scored for numbers, 2, 4, 7, 10, 11, 12, 14, and 15. These were summed to obtain a total score for the TEQ from 0 to 64. In general, higher scores indicate high levels of self-reported empathy. Scores below 45 indicate empathy levels below average. Spreng et al. (2009) reported reliability evidence for the TEQ in a sample of 200 undergraduate psychology students ($\alpha = 0.85$) and in a sample of 79 undergraduate psychology students ($\alpha = 0.85$). The TEQ has also demonstrated good test–retest reliability for scores ($r = 0.81$, $p < 0.001$) (Spreng et al., 2009).

Modified Modern Racism Scale (MMRS) (McConahay, 1986)

Participants also completed the MMRS, a modified version of the Modern Racism Scale. The MMRS differed from the original in that the words "ethnic minorities" was replaced by "Blacks". This change was done because the term "ethnic minorities" includes other minority groups not considered in this study. The MMRS consisted of 14 statements that targeted levels of explicit bias against Blacks (e.g., *"Discrimination against Blacks is no longer a problem in the United States"*). Participants were asked to indicate their level of agreement with each statement on a 7-point scale from 1, Strongly agree to 7, Strongly disagree. The MRS has satisfactory internal consistency, with coefficient alphas ranging from 0.75 to 0.86 across a large number of samples. Review of the instrument indicates stability over short periods of time, with test-retest stability results ranging from 0.72 to 0.93 across numerous samples (Sabnani & Ponterotto, 1992).

The Valuing Questionnaire (VQ) (Smout et al., 2014)

The VQ is a 10-item self-report subscale. The scale consisted of a 7-point Likert scale that ranged from 0, Not at all true to 6, Completely True. The VQ measures the "valuing" choices or actions each participant took to live consistently with their self-reported values. Only the Obstruction section of the study was used. Obstruction was defined as the disruption of valued living most likely due to experiential avoidance. For example, *"Difficult thoughts, feelings, or memories got in the way of what I really wanted to do"*. The VQ has demonstrated acceptable reliability, with internal-consistency scores of 0.89 in college student samples (Levin et al., 2017).

Likert scales

Participants were also required to complete a 13-point Likert scale that mirrored each of the six IRAP target stimuli (Barnes-Holmes et al., 2010). A total of six target stimuli were rated (i.e., three positive and three negative). The scale for each ranged from –6, Extremely positive, to +6, Extremely negative. To ensure the Likert scale served as an explicit measure of racial stereotyping, the average ratings recorded for black targets was subtracted from the average ratings recorded for White targets in each participant's scores. In other words, a positive score indicated pro-black bias (White more negative than Black), and a negative score indicated pro-White bias (Black more negative than White).

Response Latency

This variable was defined as the time (in milliseconds) that elapsed between the onset of the opening of the trial and the correct response made by the participant in the IRAP software (Hussey et al., 2015).

Procedure

This study used a repeated measures design with a control group. During the pretest and posttest phases all participants in both groups completed the TEQ, MMRS, VQ-O and Likert scale. Only the experimental group received the values and perspective taking interventions.

Familiarization IRAP

Participants accessed the IRAP electronically through their OneDrive. However, the first author was present in a remote synchronous mode to help participants to begin the test. Written instructions were provided to help participants to open the program. These described that each trial type would have one of two labels that would appear on the top of

the screen. Participants were also told that the response options, True or False would appear at the bottom of the screen and they would be required to choose an option by using either the "d" key (for True) or "k" key (for False). Both groups completed a familiarization IRAP with neutral stimuli (i.e., Insects and Flowers). This familiarization IRAP was only used during the pretest condition.

Experimental IRAP

Similarly, both groups completed the same IRAP. Instructions for the experimental IRAP were the same as the familiarization IRAP. The experimental IRAP began with a minimum of two practice blocks and four test blocks. Each block contained the same 12 trials, comprised of four different trial types. The six target stimuli included both positive and negative attributes. Three were positive (i.e., reliable, careful, and honest) and three were negative (unreliable, careless, and dishonest). The first block of the IRAP to be displayed was consistent with pro-White stereotyping (e.g., White People-reliable-True; Black People-reliable-False; White People-unreliable-False; Black People-unreliable-True) (Figure 1). The feedback presented to participants supported contingencies that alternated between pro-White and pro-black bias between each block. For example, the correct responses for pro-black bias block may be as follows: White People-unreliable-True; Black People-unreliable-False. White People-reliable-False; Black People-reliable-True. Before each block began, the participants were informed that previously correct and incorrect answers would be reversed. Trials were displayed quasi-randomly so that none of the four trial types was presented twice in a row.
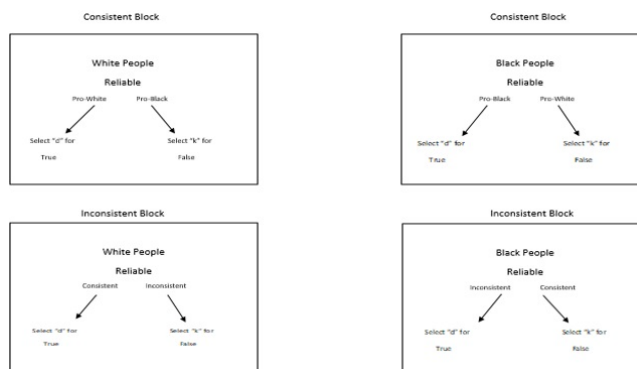

Figure 1.
A Schematic Representation of the Four IRAP Trial Types.

The instructions informed participants that the first two blocks were practice trials and errors were expected. For participants to continue on the test trials, the standard response accuracy level and response time was 80 % correct responses, median response time of 3000 ms. These criteria ensured that participants understood and were complying with the IRAP instructions. If participants failed to achieve either of the two criteria on either of the two practice blocks, their current accuracy and speed were displayed to them on the screen and they would had two additional opportunities to meet the criteria. Participants who failed to meet the criteria after six blocks would have been debriefed and their data excluded. Two participants from the experimental group failed to meet these criteria. The remaining nineteen participants proceeded to the test blocks. The procedure for the test blocks contained feedback reminding them to proceed with speed. For example, the first test block began with the instructions "test" and a reminder to "go fast".

Intervention

Experimental Group

The intervention phase was completed with the researcher in a virtual format at least one week after baseline. The components of the intervention package included values

consistency exercises and perspective taking tasks that were delivered to participants via their OneDrive.

Values Consistency Exercises. The values consistency exercises were designed to help participants clarify their values and live consistently with them as an alternative to acting upon racial bias. The first exercise used was the "Values Card Sort" (LeJeune & Luoma, 2019). This activity took approximately 5 minutes to complete. Directions for this computerized task asked participants to sort 20 value cards (e.g., Justice, Responsibility, Health) into three piles. The piles were labeled, "very important to me", "important to me", and "less important to me". The second values consistency exercise was called, "Writing an Autobiography" (Stoddard & Afari, 2014). The exercise contained two parts. The first one helped participants to identify ways they would like to ideally live their lives, whereas the second part looked at ways in which they were currently living. The general goal was highlighting that the discrepancies between the two provides a direction for values work. This exercise took approximately 5 minutes to complete.

Perspective Taking Tasks. The perspective taking tasks were displayed to participants within 10-15 minutes of beginning the values consistency tasks. The first exercise used was "The Anthropologist Metaphor" (Stoddard & Afari, 2014). The goal of this exercise was to help the participant to separate her experience from who is observing her experience. Participants were told that they could understand perspective taking by imagining their own thoughts, feelings, sensations, memories, and roles as an anthropologist might do. The second exercise was a writing task in which participants were displayed the picture of a Black man looking out a window. They were asked to write about a day in the life of the Black man pictured on the page from his perspective. The goal of the task was to determine if participants could consider how someone else might think or feel (Todd et al., 2011). This activity took approximately 5 minutes to complete.

Control Group

The control group also completed guided tasks. The first was watching a TED talk (https://www.ted.com/talks/chieh_huang_how_to_know_if_it_s_time_to_change_careers?language=en). They were then asked to complete two short writing tasks about the best job they have ever held and their ideal job. Each task took approximately 5 minutes to complete. Participants were then asked to complete the same perspective taking writing task as the experimental group; however, they were not told from what perspective to write it.

Social Validity Measure

Participants completed a post-study survey. The form was a Likert-style questionnaire. Answers ranged from one to five with 1 meaning, Extremely or Definitely and 5, Not Easy or Definitely Not. For example, "How easy was it to understand the difference between values and goals?", "Are you aware of any personal racial biases?", and "Did you feel your participation in the process affected any biases you may have had?" The sum of three items created a total score. An open-ended comment section was also provided. Participants returned these questionnaires via OneDrive during the last week of the intervention.

Analytic Plan

A number of statistical analyses were conducted in line with previous research using the IRAP (Hussey et al., 2015). A 2 x 4 mixed repeated-measures analysis of variance ANOVA was conducted to determine if there was interaction effect between group and trial types. In addition, a post-hoc test with Bonferroni adjustment was conducted. Bonferroni adjusted levels were calculated by dividing alpha level 0.05 by the number of post-hoc tests involved ($\alpha < 0.0125$), for pairwise comparisons and independent samples t-tests. Four independent

sample . tests also with Bonferroni adjustment (adjusted level = .00125) were conducted for the four trial types for both groups. For the explicit measures a one-way multivariate analysis of variance MANOVA was conducted with group as the independent variable and the explicit measures as dependent variables. An independent t test was conducted to explore the differences in the dependent variables between control and experimental groups. An alpha level of 0.025 was used. Finally, a correlation matrix was created to check for relationships between the implicit and explicit measures.

Results

Response Latencies

To control for the individual variations in speed of responding that could be confounding when analyzing between-group differences, each participant's response latency data was transformed into DIRAP scores (Hussey et al., 2015). The response latencies for the mean DIRAP scores for the four trial types (i.e., Consistent-White, Inconsistent-White, Consistent-Black, Inconsistent-Black) are shown in Figure 2. The data showed that responses for the Consistent-White, Inconsistent-White, and Inconsistent Black trial types were in the predicted direction for both groups. Both groups responded "True" more quickly than "False" for the Consistent-White questions (reliable-White, honest-White, careful-white) as compared to the Consistent-Black (unreliable-Black, dishonest-Black, careless-Black). The effect for the Consistent-Black trial type was in a non-predicted direction for both groups, with a stronger effect in the experimental group.
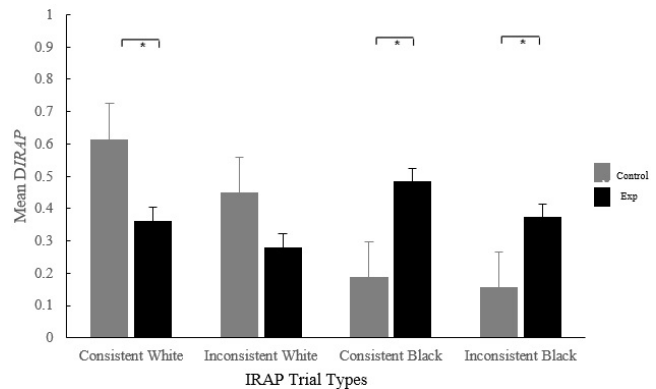

Figure 2.
D-IRAP Trial-type Scores after Trials 3 and 4 were inverted.

A 2 x 4 mixed repeated-measures analysis of variance ANOVA was conducted to determine if there was interaction effect between group and trial types. This analysis revealed a significant main effect for group type F (1, 148) = 25.37, p < 0.001, n_p^2 = 0.146) and trial type, F (3, 148) = 20.5, p ≈ 0, n_p^2 = 0.293), but no interaction effect between group and trial type for context or interaction (p = 0.613). A Bonferroni multiple comparisons test revealed a significant difference between Consistent-White and Inconsistent-White trial types, (p < 0.000), Consistent-White and Inconsistent-Black trial types, (p < 0.000), Consistent-Black and Inconsistent-White trial types, (p ≈ 0), and Inconsistent-White and Inconsistent-Black trial types (p = 0.005). Four independent samples t test were run to determine if there was a significant difference in the means between groups for trial type. The test revealed a significant difference for Consistent-White (t = -3.23, df = 37, p = 0.003), Consistent-Black (t = -2.03, df = 37, p = 0.049), and Inconsistent-Black trial types (t = -3.240, df = 37, p = 0.003). No significant effect was found for the Inconsistent-White trial type (t = -1.708, df = 37, p = 0.095).

Internal reliability

To calculate reliability for the IRAP, two overall *D* scores following the same steps as for the DIRAP

scores above. The split-half correlations, applying Spearman-Brown corrections, between odd and even DIRAP scores, calculated separately for the two groups, proved to be weak and insignificant for the control group, $r = 0.34$, $n = 20$, (Cronbach's alpha = 0.57), and the experimental group $r = 0.34$, $n = 19$, (Cronbach's alpha = 0.61).

Explicit Measures

To determine if there were any differences between groups on more than one continuous dependent variable, a one-way multivariate analysis of variance MANOVA was conducted with group as the independent variable and the explicit measures as dependent variables. (i.e., MMRS, TEQ, VQ, Likert, Table 2). There was a statistically significant difference between experimental and control groups when considered jointly along the dependent variables. $F(4, 34) = 3.73$, $p = 0.013$, Wilk's $\Lambda = 0.695$, $n^2 = 0.305$.

Table 2
Mean and Standard Deviation for the Study Variables

Groups	TEQ				MMRS				VQ			
	Pretest		Posttest		Pretest		Posttest		Pretest		Posttest	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Experimental	35.7	2.75	35.9	2.41	31.1	2.01	33	2.10	7.94	4.74	8.37	5.39
Control	36	2.36	37.1	2.48	31.1	3.12	31.5	3.20	2.8	2.04	2.8	2.04

Note.*TEQ* = Toronto Empathy Questionnaire;
MMRS = Modified Modern Racism
Scale; *VQ* = Valuing Questionnaire.

An independent t test was conducted to explore the differences in the dependent variables between control and experimental groups. An alpha level of 0.025 was used. For the MMRS, both groups were normally distributed. Variances were homogeneous, $F(1, 37)$, $p = 0.093$. No significant difference was evident between control and experimental groups, $t(37) = 1.717$, $p = 0.093$. A medium effect size was noted, $\Delta = 0.711$. For the VQ, both groups were normally distributed, but variances were not homogeneous $F(1, 37) = 17.085$, $p < 0.001$. Thus, equal variance was not assumed.

A statistically significant difference was evident between the experimental and control groups, $t(22.826) = 4.219$, $p < 0.001$. A large effect size was noted, $\Delta = 1.031$. For the TEQ, both groups were normally distributed, and variances were homogeneous. $F(1,37) \approx 0$, $p = 0.991$. No significant difference was evident between experimental and control group, $t(37) = -1.405$, $p = 0.168$. A negligible effect size was noted, $d \approx 0$. For the Likert survey, both groups were normally distributed, and variances were homogeneous, $F(1,37) = 1.427$, $p = 0.240$. No significant difference was evident between the experimental and control groups, $t(37) = -0.889$, $p = 0.380$. A small effect size was noted 0.31.

Implicit-Explicit Correlations

To determine if a significant relationship existed between the implicit and explicit measures, two correlation matrices for the implicit and explicit measures and groups were calculated: one for the control and experimental conditions. Each matrix correlated the four trial-types and overall DIRAP scores with each of the explicit measures. All correlations between the DIRAP scores and four explicit measures proved to be nonsignificant ($p > 0.05$).

Social Validity

The mean for Question 1, "How easy was it to understand the difference between values and goals?" was ($M = 3.42$). The mean for Question 2, "The intervention had a significant impact on my ability to identify my values?" was ($M = 2.74$). The mean for Question 3, "This intervention helped me to make decisions based on my values" was ($M = 2.79$). Overall, the results indicated that the values clarification activities used with the experimental group were somewhat easy to complete, the intervention may have had a significant impact on their ability to identify their values, and the intervention may have helped them learn to make decisions based on their values.

Discussion

The current study investigated the effects of two interventions, values clarification and perspective taking, using the IRAP and multiple self-report measures on racism, empathy, and values. The findings of the current study provided additional support for the IRAP as a behavior-based implicit measure of racial bias (Barnes-Holmes et al., 2010; Drake et al., 2015; Dunne et al., 2018; Mizael & de Almeida, 2019; Power et al., 2017a, 2017b). Overall, after the values work and perspective taking exercises, participants in the experimental group recorded shorter mean responses for Inconsistent-Black trial blocks versus Consistent-Black trial blocks. These results indicated that participants responded faster when the correct response was to affirm positive-Black and slower when the correct response was to affirm negative-Black attributes. In contrast, the control group recorded shorter mean response latencies for Consistent-Black versus Inconsistent Black trial-blocks. These results indicated that participants responded faster when the correct response was to affirm positive-White and negative-Black attributes. These results are consistent with the body of research on the positive effects of values (Arauz et al., 2017; Graham et al., 2015; Lillis & Hayes, 2007; West et al., 2013), and perspective taking (Pashak et al., 2018; Todd et al., 2011; Wang et al., 2014) in decreasing racial bias.

One hypothesis for these results is that the intervention package encouraged a mindful awareness of the experimental participants' own prejudicial bias, which in turn facilitated the ability to respond in line with their personal values. Unlike their control counterparts, experimental participants were asked to sort their values and then complete a behavioral expression of their values (i.e., perspective taking writing task). It seems plausible therefore, that both the tasks and their order contributed to this outcome. Support for this hypothesis can be found in earlier research by Lillis and Hayes (2007), which suggests that the probability of behaving in a values consistent manner is enhanced by experiential exercises.

Additional support may be found in Todd et al. (2011). These researchers argued that individuals can "experience inner conflict" because private events (e.g., thoughts, feelings) conflict with their values (p. 2). In other words, an individual can behave in a manner inconsistent with a personal disavowal of prejudice. This might explain why both groups self-reported lower levels of explicit bias on questionnaires as compared to those recorded by the IRAP. These finding appear to suggest that an ACT intervention package is warranted in the development of a more effective technology to decrease racial bias (Banks et al., 2021; Garcia et al., 2021; Matsuda et al., 2020; Williams et al., 2020).

A significant difference between Consistent-White and Inconsistent-White trial types, ($p \approx 0$), Consistent-White and Inconsistent-Black trial types, ($p \approx 0$), Consistent-Black and Inconsistent-White trial types, ($p \approx 0$), and Inconsistent-White and Inconsistent-Black trial types ($p = 0.005$) between experimental and control groups was found. However, Inconsistent-Black was the only trial-type across all attributes (Reliable, Careful, Honest) for which opposing bias between the experimental and control groups was consistently demonstrated. That is, the former demonstrated pro-black bias while the latter demonstrated pro-White bias. Consistent-Black was the only trial type for which responses for both groups across all attributes (Reliable, Careful, Honest) was consistently pro-Black. This last result was unexpected among the control group, especially as compared to the results of the Inconsistent-Black trial type.

Considering that no incentives were offered for participation, one explanation may be that those who volunteered for the study did so because they knew the purpose of the study was to better understand how racial bias is formed and the act of participating was consistent with a personal value of equality or justice. This might explain why mean control group scores were recorded as pro-Black despite there was no exposure to values clarification and perspective-taking intervention. When the results for the

Consistent-Black trial type are analyzed in conjunction with those of the Inconsistent-Black trial type, the data suggest that participants reflected a pre-experimental learning history of reinforcement for relating positive attributes with both White and Black persons. However, with regard to the Inconsistent-Black trial type, the results suggest that the experimental group responses were affected by the values clarification and perspective taking intervention (Cohen et al., 2006; Todd & Simpson, 2017; West et al., 2013).

Additionally, a statistically significant difference was discovered for VQ scores, $p < 0.001$ between the control and experimental groups. Interestingly, the experimental group recorded higher scores (M = 8.36) as compared with the control group (M = 2.8). This was an unexpected result as higher self-reported scores represent more obstruction, or interference with living consistently with one's values and this group scored lower across all pro-White trial types. It is hypothesized that after receiving the intervention, experimental group participants self-reported a more factual frequency of aversive private events that formed barriers to living in line with their values as compared to control group participants.

Although no significant correlations between the explicit and implicit measures were found, significant correlations between group and trial type were recorded suggesting that identifying personal values in coordination with completion of a brief perspective taking task may inhibit implicit racial bias. These findings are supported by previous studies that demonstrated the effectiveness of values consistency (Arauz et al., 2017; Banks et al., 2021; Miller et al., 2020; West et al., 2013). These studies reported that even a brief values clarification intervention may mitigate self-reported experiences of distress and negative affect (West et al., 2013) and a strong correlation between acceptance and lower levels of distress for participants who reported higher levels of acceptance and values consistency. In addition, Banks et al. (2021) reported that an ACT intervention decreased depression, anxiety, and internalized shame, while Miller et al. (2020)

found that psychological flexibility mitigated the effect of racial stressors faced by minority students.

In sum, as hypothesized, participants in the experimental group demonstrated a decreased level of pro-White bias after exposure to the values clarification and perspective taking intervention package thus supporting previous research that perspective taking, and values consistency are promising interventions for racism (Banks et al., 2021; Cohen et al., 2006; Davis et al., 2021; Todd et al., 2011; Todd et al., 2012).

Limitations

Some limitations were noted in the current study. First, the IRAP software used was written in JAVA and is not compatible with devices that run iOS or Chrome OS. As a result, many prospective participants were rejected on the basis of their computer's operating system. Future researchers might consider working with or getting permission from the original developers to create a program that is compatible with Apple and Chrome devices. Second, the study was completed remotely, and while the IRAP software is relatively easy to install, it also requires the installation of a separate stimuli file. While the researcher did provide a task analysis for installation and was present via video call to guide the participants through the process, many participants that lacked advanced computer skills experienced frustration during the set-up process. Future researchers could eliminate these issues by providing a physical location for participants with a preprogrammed computer. Geist et al. (2023) have developed a Web-IRAP that can be used for future IRAP studies conducted via online. Third, approval and recruitment for the study were extremely difficult. This is most likely due to the current racial tensions in the United States and might explain the general reluctance of participants to volunteer. Future researchers intending to accumulate data on implicit bias within the United States will need to identify and secure

various sources of recruitment. Fourth, the current study did not include a generalization probe. Future research could include a follow-up IRAP to determine the long-term effects of the values consistency and perspective taking. Finally, the current study was completed with a very small sample of participants who identified as White. In addition, an inherent limitation of self-identification was that participant identity was not verified. This may have introduced a confusion. When considering the results for the Consistent-Black trial type, for example, as the only trial type for which responses across all attributes were consistently pro-black, further exclusionary criteria may have functioned to expose the effects of exposure to diversity. However, questions regarding family of origin and race were excluded during the IRB process. Therefore, the results of this study cannot be said to represent the population at large.

Future Research

An initial power analysis suggested that the ideal participant size would be almost twice the size of this study. It is possible that as a result of smaller sample, the statistical analyses conducted did not reveal the true effect of the intervention. Future research should also extend the duration of exposure to values and perspective taking exercises. Furthermore, adding other ACT components, for example, committed actions, might encourage participants to take measurable steps to decrease racial bias. Future research should be more inclusive. For example, additional control and experimental groups consisting of persons who identify as minorities could be added and their data could be compared to those of White participant groups. If we concede that implicit bias is the result of relational responding, then future research should also include groups that are separated along generational, political, geographical, socio-economic, and educational variables. Understanding how these variables affect implicit bias as environmental factors, might help behavior analysts in the applied setting to develop more appropriate

interventions. Future research should also include children to help identify if there is an optimal age at which implicit racial bias can be prevented.

Conclusions

The goals of the current study were to determine if a joint intervention of perspective taking, and values consistency would be effective for decreasing racial bias. Based on the results of the current study, the experimental group participants self-reported much higher levels of obstruction with regard to living their lives in line with their values. However, response latency data for the IRAP showed that this group scored higher with regard to pro-Black bias. While the exact reasons are unknown, we can hypothesize that it is that after completing the exercises on values consistency and perspective taking, that participants observed that they were not living in line with their values and committed to actions (Banks et al., 2021; Cohen et al., 2006; Lillis & Hayes, 2007; Williams et al., 2020). Matsuda et al. (2020) argued that biases are often difficult to change because they represent a lengthy learning history that often interferes with an individual's ability to form new relations. Based on the results of the current study, an intervention package that includes even limited exposure to a portion of values consistency and perspective taking skills can have an effect on implicit racial bias (Cohen et al., 2006; Davis et al., 2021; Kanter et al., 2020; Lillis & Hayes, 2007; West et al., 2017).

References

Albright, L. K., Blair, B. J., & Ferman, D. M. (2021). Stimulus equivalence: A derived relational account of cultural biases. In J. Sadavoy & M. Zube (Eds). *A scientific framework for compassion and social justice: Lessons in applied behavior analysis* (pp. 291-296). Taylor & Francis/Routledge.

Arauz, J., Danitz, S. B., Orsillo, S. M., & Coyne, L. W. (2017). A preliminary exploration of education values, distress, and acceptance

among self-identified white and non-white incoming college freshmen at a private university. *Journal of Contextual Behavioral Science, 6*(3), 288-292. https://doi.org/10.1016/j.jcbs.2017.05.001

Banks, K. H., Goswami, S., Goodwin, D., Petty, J., Bell, V., & Musa, I. (2021). Interrupting internalized racial oppression: A community based ACT intervention. *Journal of Contextual Behavioral Science, 20*, 89-93. https://doi.org/10.1016/j.jcbs.2021.02.006

Barnes-Holmes, D., Harte, C., & McEnteggart, C. (2020). Implicit cognition and social behavior. In R. A. Rehfeldt, J. Tarbox, & M. Fryling (Eds.), *Applied Behavior Analysis of Language and Cognition* (pp. 264-280). New Harbinger.

Barnes-Holmes, D., Murphy, A., Barnes-Holmes, Y., & Stewart, I. (2010). The Implicit Relational Assessment Procedure: 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*, 57-79. https://doi.org/10.1007/BF03395694

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science, 313*(5791), 1307-1310. https://doi.org/10.1126/science.1128317

Davis, C. H., Krafft, J., Hicks, E. T., & Levin, M. E. (2021). The role of psychological inflexibility and perspective taking in anti-racism and anti-sexism. *Personality and Individual Differences, 175*, 110724. https://doi.org/10.1016/j.paid.2021.110724

de Carvalho, M. P., & de Rose, J. C. (2014). Understanding racial attitudes through the stimulus equivalence paradigm. *The Psychological Record, 64*, 527-536. https://doi.org/10.1007/s40732-014-0049-4

Dixon, M. R., Branon, A., Nastally, B. L., & Mui, N. (2009). Examining prejudice towards Middle Eastern persons via a transformation of stimulus functions. *The*

Behavior Analyst Today, 10(2), 295-318. https://doi.org/10.1037/h0100672

Dixon, M. R., & Hayes, S. C. (2022). On the disruptive effects of behavior analysis on behavior analysis: The high cost of keeping out acceptance and commitment therapy and training. *Behavior Analysis in Practice.* https://doi.org/10.1007/s40617-022-00742-4

Dixon, M. R., & Lemke, M. (2007). Reducing prejudice towards Middle Eastern persons as terrorists. *European Journal of Behavior Analysis, 8*(1), 5-12. https://doi.org/10.1080/15021149.2007.11434269

Drake, C. E., Kramer, S., Sain, T., Swiatek, R., Kohn, K., & Murphy, M. (2015). Exploring the reliability and convergent validity of implicit racial evaluations. *Behavior and Social Issues, 24*, 68-87. https://doi.org/10.5210/bsi.v24i0.5496

Dunne, C., McEnteggart, C., Harte, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2018). Faking a race IRAP effect in the context of single versus multiple label stimuli. *International Journal of Psychology and Psychological Therapy, 18*(3), 289-300. https://www.ijpsy.com/volumen18/num3/496.html

Ferroni-Bast, D., Fitzpatrick, J., Stewart, I., & Goyos, C. (2019). Using the Implicit Relational Assessment Procedure (IRAP) as a measure of reaction to perceived failure and the effects of a defusion intervention in this context. *The Psychological Record, 69*, 551-563. https://doi.org/10.1007/s40732-019-00349-2

FitzGerald, C., Martin, A., Berner, D., & Hurst, S. (2019). Interventions designed to reduce implicit prejudices and implicit stereotypes in real world contexts: A systematic review. *BMC psychology, 7*(29), 1-12. https://doi.org/10.1186/s40359-019-0299-7

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology, 117*(3), 522-559. https://doi.org/10.1037/pspa0000160

Fryling, M., Rehfeldt, R. A., Tarbox, J., & Hayes, L. J. (Eds.). (2020). *Applied behavior analysis of language and cognition: Core concepts and principles for practitioners*. New Harbinger Publications.

Garcia, Y., Andrews, M., & Brothers, L. (2021). A relational frame approach to understanding perspective-taking in compassion and social justice. In J. Sadavoy & M. Zube (Eds). *A scientific framework for compassion and social justice: Lessons in applied behavior analysis* (pp. 217-224). Taylor & Francis/Routledge.

Graham, J. R., West, L., & Roemer, L. (2015). A preliminary exploration of the moderating role of valued living in the relationships between racist experiences and anxious and depressive symptoms. *Journal of Contextual Behavioral Science, 4*(1), 48-55. https://doi.org/10.1016/j.jcbs.2014.11.001

Geist, T., Plezia, S., Cepeda-Benito, A., López, M. H., & Valverde, M. R. (2023). Online administration of the Implicit Relational Assessment Procedure: The Web-IRAP. *The Psychological Record.* https://doi.org/10.1007/s40732-022-00533-x

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational Frame Theory: A Post-Skinnerian account of human language and cognition*. Plenum Press.

Hayes, S. C., Strosahl, K., & Wilson, K. (2011). *Acceptance and commitment therapy: The process and practice of mindful change*. Guilford Press.

Hussey, I., Thompson, M., McEnteggart, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2015). Interpreting and inverting with less cursing: A guide to interpreting IRAP data. *Journal of Contextual Behavioral Science, 4*(3), 157-162. https://doi.org/10.1016/j.jcbs.2015.05.001

Kanter, J. W., Williams, M. T., Kuczynski, A. M., Corey, M. D., Parigoris, R. M., Carey, C. M., Manbeck, K. E., Wallace, E. C., & Rosen, D. C. (2020). The measurement and structure of microaggressive communications by white people against black people. *Race and Social Problems, 12*, 323-343. https://doi.org/10.1007/s12552-020-09298-w

Kishita, N., Muto, T., Ohtsuki, T., & Barnes-Holmes, D. (2014). Measuring the effect of cognitive defusion using the implicit relational assessment procedure: An experimental analysis with a highly socially anxious sample. *Journal of Contextual Behavioral Science, 3*(1), 8-15. https://doi.org/10.1016/j.jcbs.2013.12.001

LeJeune, J. & Luoma, J. B. (2019). *Values in therapy: A clinician's guide to helping clients explore values, increase psychological flexibility, and live a more meaningful life*. New Harbinger Publications.

Levin, M. E., Haeger, J. A., Pierce, B. G., & Twohig, M. P. (2017). Web-based acceptance and commitment therapy for mental health problems in college students: A randomized controlled trial. *Behavior Modification, 41*(1), 141–162. https://doi.org/10.1177/0145445516659645

Lillis, J., & Hayes, S. C. (2007). Applying acceptance, mindfulness, and values to the reduction of prejudice: A pilot study. *Behavior Modification, 31*(4), 389-411. https://doi.org/10.1177/0145445506298413

Matsuda, K., Garcia, Y., Catagnus, R., & Brandt, J. A. (2020). Can behavior analysis help us understand and reduce racism? A review of the current literature. *Behavior Analysis in Practice, 13*, 336-347. https://doi.org/10.1007/s40617-020-00411-4

Mcconahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91–125). Academic Press.

Mizael, T. M., & de Almeida, J. H. (2019). Revisão de estudos do implicit relational assessment procedure sobre vieses raciais. *Acta Comportamentalia: Revista Latina de Análisis del Comportamiento*, 27(4), 437-461. https://www.revistas.unam.mx/index.php/acom/article/view/72025

Mizael, T. M., & Sampaio, A. A. (2019). Racismo institucional: Aspectos comportamentais e culturais da abordagem policial.

Acta Comportamentalia: Revista Latina de Análisis del Comportamiento, 27(2), 215-231. https://www.revistas.unam.mx/index.php/acom/article/view/69861

Mizael, T. M., de Almeida, J. H., Silveira, C. C., & de Rose, J. C. (2016). Changing racial bias by transfer of functions in equivalence classes. *The Psychological Record, 66,* 451–462. https://doi.org/10.1007/s40732-016-0185-0

Obasogie, O. (2013). *Blinded by sight*. Stanford University Press.

Pashak, T. J., Conley, M. A., Whitney, D. J., Oswald, S. R., Heckroth, S. G., & Schumacher, E. M. (2018). Empathy diminishes prejudice: Active perspective-taking, regardless of target and mortality salience, decreases implicit racial bias. *Psychology, 9*(6), 1340-1356. https://doi.org/10.4236/psych.2018.96082

Power, P. M., Harte, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2017a). Exploring racial bias in a European country with a recent history of immigration of Black Africans. *The Psychological Record, 67,* 365-375. https://doi.org/10.1007/s40732-017-0223-6

Power, P. M., Harte, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2017b). Combining the implicit relational assessment procedure and the recording of event related potentials in the analysis of racial bias: A preliminary study. *The Psychological Record, 67,* 499-506. https://doi.org/10.1007/s40732-017-0252-1

Roberts, S. O. & Rizzo, M. T. (2021). The psychology of American racism. *American Psychologist, 76*(3), 475-487. https://doi.org/10.1037/amp0000642

Sabnani, H. B. & Ponterotto, J. G. (1992). Racial/ethnic minority-specific instrumentation in counseling research: A review, critique, and recommendations. *Measurement & Evaluation in Counseling and Development, 24*(4), 161-187. https://psycnet.apa.org/record/1992-21983-001

Saini, V. & Vance, H. (2020). Systemic racism and cultural selection: A preliminary analysis of metacontingencies. *Behavior and Social Issues, 29,* 52-63. https://doi.org/10.1007/s42822-020-00040-0

Salter, P. S., Adams, G., & Perez, M. J. (2018). Racism in the structure of everyday worlds: A cultural-psychological perspective. *Current Directions in Psychological Science, 27*(3), 150-155. https://doi.org/10.1177/0963721417724239

Smout, M., Davies, M., Burns, N., & Christie, A. (2014). Development of the Valuing Questionnaire (VQ). *Journal of Contextual Behavioral Science, 3*(3), 164-172. https://doi.org/10.1016/j.jcbs.2014.06.001

Spreng, R. N., McKinnon, M. C., Mar, R. A., & Levine, B. (2009). The Toronto Empathy Questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of Personality Assessment, 91*(1), 62-71. https://doi.org/10.1080/00223890802484381

Stoddard, J. A., & Afari, N. (2014). *The big book of act metaphors: A practitioner's guide to experiential exercises and metaphors in acceptance and commitment therapy*. New Harbinger Publications.

Tarbox, J., Szabo, T. G., & Aclan, M. (2022). Acceptance and commitment training within the scope of practice of applied behavior analysis. *Behavior Analysis in Practice, 15*(1), 11-32. https://doi.org/10.1007/s40617-020-00466-3

Todd, A. R., Bodenhausen, G. V., & Galinsky, A. D. (2012). Perspective taking combats the denial of intergroup discrimination. *Journal of Experimental Social Psychology, 48*(3), 738-745. https://doi.org/10.1016/j.jesp.2011.12.011

Todd, A., R., Bodenhausen, G. V., Richeson, J. A., & Galinsky, A. D. (2011). Perspective taking combats automatic expressions of racial bias. *Journal of Personality and Social Psychology, 100*(6), 1027–1042. https://doi.org/10.1037/a0022308

Vahey, N. A., Nicholson, E., & Barnes-Holmes, D. (2015). A meta-analysis of criterion effects for the Implicit Relational

Assessment Procedure (IRAP) in the clinical domain. *Journal of Behavior Therapy and Experimental Psychiatry, 48,* 59-65. https://doi.org/10.1016/j.jbtep.2015.01.004

Wang, C. S., Kenneth, T., Ku, G., & Galinsky, A. D. (2014). Perspective-taking increases willingness to engage in intergroup contact. *PloS one, 9,* e85681. https://doi.org/10.1371/journal.pone.0085681

Watt, A., Keenan, M., Barnes, D., & Cairns, E. (1991). Social categorization and stimulus equivalence. *The Psychological Record, 41*(1), 33–50. https://doi.org/10.1007/BF03395092

Williams, M. T., Kanter, J. W., Peña, A., Ching, T. H., & Oshin, L. (2020). Reducing microaggressions and promoting interracial connection: The racial harmony workshop. *Journal of Contextual Behavioral Science, 16,* 153-161. https://doi.org/10.1016/j.jcbs.2020.04.008

Notes