Functional Differences in the Words the Japanese use for "I" in Relation to Self-Esteem: Key Factor for Establishing Level of Self-Esteem in Psychoanalysis

Pin ZHANG*[1], Shinji TANI*[2]

Abstract:

The current study investigated the functional differences of "own name" depending on the level of self-esteem in the Implicit Relational Assessment Procedure (IRAP) by observing the results of one-sample t-tests and ANOVAs of two higher and lower self-esteem groups. According to the suggestion, the lower the self-esteem, the weaker the trial-type 1 IRAP effect measured by the name IRAP by Zhang and Tani (2022). We assumed that the single-trial-type-dominance effect (STTDE) appears only in the group with high self-esteem and is not observed in the group with low self-esteem for IRAP. Participants were thirty-one native Japanese speakers, and for the IRAP, two labels (own name or others) and two targets (positive or negative words) were used. Each participant completed the Japanese version of the Rosenberg Self-Esteem Scale (RSES-J) and IRAP. We divided the participants into higher and lower self-esteem groups based on the results of the RSES-J.

Results found a significant difference for the participant's name-positive trial type (t = 6.14, df = 15, $p < .01$) for the higher self-esteem group. One-way repeated-measures ANOVAs yielded a significant main effect of trial-type on the IRAP for the higher self-esteem group (F (3, 15) = 10.35, $p < .01$, $\eta\rho 2 = .41$). Bonferroni post hoc tests indicated that the effects for the participant's name-positive differed significantly from the three other trial-types ($p < .01$); the remaining types did not differ significantly from each other ($p > .05$) for the higher self-esteem group. For the lower self-esteem group, ANOVAs confirmed no significant main effect (F (3, 14) = 1.30, $p > .05$). Bonferroni post hoc tests indicated that the effect for all trials did not differ significantly from each other ($p > .05$).

The hypothesis was supported by the results that STTDE appears only in the group with high self-esteem and is not observed in the group with low self-esteem for IRAP. In the results of a differential arbitrarily applicable relational responding effects analysis of word functions, the function of "own name" differed between respondents with low versus high self-esteem.

---

*1 Doctoral Student, Graduate School of Human Science, Ritsumeikan University

*2 Professor, College of Comprehensive Psychology, Ritsumeikan University

Email: *1 gr0316ei@ed.ritsumei.ac.jp

*2 stt23197@pl.ritsumei.ac.jp

Received on 2023/1/27, accepted after peer reviews on 2023/4/14.

Journal of the Asia-Japan Research Institute of Ritsumeikan University, 2023.
PRINT ISSN 2435-0184 ONLINE ISSN 2435-0192, Vol.5, pp.1-15.

In conclusion, the study successfully identified the differences in that the higher the self-esteem the stronger was the appetitive evoking function, and the lower the self-esteem the stronger the aversive evoking function for "own name" in Name-IRAP.

Keywords: *Implicit Relational Assessment Procedure, self-esteem, Differential Arbitrarily Applicable Relational Responding Effects, single-trial-type dominance effect, evoking orienting relating*

1. Introduction

The self-concept is one of the important focuses of clinical psychology. For instance, from the perspective of the person-centered approach, Feltham and Horton (2006) noted that "the self-concept is a central aspect of the person-centered approach to counseling. It is basically how people define themselves, for example, "I am caring, I am cheerful, I can sometimes be funny." It is a fluid but consistent pattern of perceptions of the "I" or "me" in relation to the environment, personal values, goals, and ideals. Acceptance and commitment therapy is one of the most representative and empirically supported therapies of the so-called contextual therapies (Hayes, 2004; Ruiz, 2010). ACT identifies self-concept as a factor influencing psychological flexibility skills that the ability to act based on one's own values even in the presence of difficulties and negative emotions. From the above, self-concept is an important topic for various psychotherapies; the word "I" is often used in clinical situations or interventions. However, in the case of the Japanese language, it has many words to refer to the self that mean the same as "I" in English, including *watashi, boku, ore*, and *jibun*. Hirose and Hasegawa (2016) explained that Japanese is a highly context-dependent language, especially with respect to *watashi, jibun, ore,* etc., which depends heavily on the context. The authors note that the Japanese self-meaning words refer to the social concept (e.g., interpersonal relations), which depends heavily on the contextual change of the listener. In detail, Zhang and Tani (2021) summarized that Japanese self-meaning is considered to have the following two functions: first, signifying an interpersonal relationship between a speaker and a listener; for example, the words *watashi* or *boku* are often used when the speaker talks to an elder, and the word *ore* is often used when the speaker talks to someone younger or the same age. The second function is signifying whether they are in a private or public setting. For example, the word *watashi* is often used for talking with the listener in a public setting (e.g., presenting a speech in front of many people). However, the word *boku* is used for private conversations. It is important to note that when a person uses an incorrect self-meaning word in a given context, they would be punished or corrected with a word choice by the listener in Japan. If a man says *ore* when talking with elders, he would be avoided as a person who does not know correct manners or would be frowned at. In addition to this, for the Japanese culture, the self-meaning words are different depending on gender. For instance, in private, "*ore*" and "*boku*" are often used by men, while "*atashi*" and "*uchi*" are often used by women. Since the function of self-meaning words is considered a necessary role for the self-concept in the clinical counseling, in order to understand the issues involved in self-concept for Japanese, it is important to clarify the functional differences of self-meaning words in the Japanese language.

However, studies that examine the functional differences between words that refer to the self to date are limited. Only one published study (Zhang and Tani, 2022) used the implicit relational

assessment procedure (IRAP; Barnes-Holmes et al., 2006) and arbitrarily applicable relational responding effects model (DAARRE; Barnes-Holmes et al., 2020) to investigate whether there are functional differences between "*watashi*" and "own name."

IRAP is a behavior-analytic measure of implied relationships useful for analyzing the functions of words that refer to the self. It is a computer program that presents label and target stimuli on a monitor and asks participants to identify the relationship (e.g., yes/no), and the outcome of interest is relationship response times that are consistent and inconsistent with the learning history (Barnes-Holmes et al., 2010b). The difference in response times between the consistent and inconsistent blocks is the IRAP effect, "assumed to provide a non-relative index of the strength of relational responding being assessed" (Hughes et al., 2011; Barnes-Holmes et al., 2017).

To set the IRAP task, the researcher should set two types of label stimuli (for example, label 1 and label 2), two types of target stimuli, and two response options. The 2×2 relationship of two label stimulus sets (e.g., "I am" and "I am not"), and two target stimulus sets (e.g., "kind" and "unkind") produces four trial-types. Trial-type 1 "label 1 and target 1 ('I am' and 'kind')," Trial-type 2, "label 1 and target 2 ('I am' and 'unkind')", Trial-type 3, "label 2 and target 1 ('I am not' and 'kind')," and Trial-type 4, "label 2 and target 2 ('I am not' and 'unkind')." On each trial, participants are required to choose one of two response options (e.g., "True" or "False"), indicating the relation between the label and target stimulus. These trial-types will appear in two types of blocks in a random display, defined as consistent and inconsistent (i.e., with or without the individual learning history of participants). Participants are required to choose response options to blocks of these stimulus pairings as accurately and quickly as possible according to two consistent or inconsistent responding rules (Hussey et al., 2015). The correct response option in the consistent block will be completely opposite to the correct response option in the inconsistent block (i.e., if the correct response option is "Yes" in the consistent blocks, the response option "No" will be correct in the inconsistent blocks). The basic IRAP hypothesis is that response times (latencies) will be faster for trials under the history consistent trials responding rule (e.g., I am kind - True) rather than in the history inconsistent trials under the other rule (e.g., I am kind - False). For the IRAP result, each trial-type produces a different IRAP effect. The greater the difference in the time of responding to a type of relationship (consistent or inconsistent), the stronger (degree of bias from 0) the IRAP effect.

DAARRE is an analysis model that analyzes the function of stimuli from the results of IRAP. The DAARRE model explains the IRAP effect related to the interactions between the functional properties of the stimuli and the relationships between the label and target stimuli, and the model organizes these effects into three categories: 1) Crel: the relationship between the two stimuli presented in each IRAP trial; 2) Cfunc: the orienting (based on the prior frequency of use) or evoking (based on the extent to which they evoke appetitive or aversive) function of the two stimuli themselves; or 3) RCIs: relational coherence indicators, that is the coherence functions presented in the response options (Barnes-Holmes et al., 2020). Briefly, people are more likely to respond to a stimulus when the relationships among the stimuli are consistent with their learning history (e.g., response in "I am human" faster than "I am not human"). However, the DAARRE explained that even when the stimulus relationships are consistent with their learning history, they are more likely to respond faster when one stimulus has a stronger orienting function than the other (e.g., higher frequency of contact indicates a stronger orienting function) and a more appetitive than aversive evoking function. The DAARRE model explains the influence of the stimulus's own function on human response time, and by using this explanation, it is possible to analyze what functional differences exist based on differences in response time to various

stimuli (differences in IRAP effects).

In order to investigate the functional differences of "*watashi*" and "own name" for self-concept, Zhang and Tani (2022) used a stimulus setting that measures self-concept (Self-IRAP, Pidgeon et al., 2020). In their "Name-IRAP," participants "own name" was label 1 and the word "others" was label 2. The target stimuli were positive (target 1, good, success, honest, capable, pleasant, and confident in Japanese) and negative words (target 2, bad, failure, dishonest, worthless, nasty, and ashamed in Japanese), and the response options were *hai* (yes in English) and *iie* (no in English). In their "I-IRAP," researchers employed the word "*watashi*" as label 1, the other stimulus was the same as "Name-IRAP." Therefore, the difference between the IRAPs is whether to use "own name" or "*watashi*" to refer to the self. If different IRAP effects are observed in trial-type 1 "*watashi* or 'own name' and positive words," and trial-type 2, "'*watashi'* or 'own name' and negative words" between the two IRAPs, this suggests that "own name" and "*watashi*" may be functionally different.

However, the results found the same IRAP effect called the single-trial-type-dominance effect (STTDE; Finn et al., 2017). The STTDE is defined as a pattern of relational response in which one of the IRAP trial types produces an effect that is significantly larger than the effects for the other three. For the study (Zhang and Tani., 2022), the two IRAPs have a same significantly stronger trial-type 1 IRAP effect than the other types. Therefore, no effective difference in IRAP (no significant difference in response time) was observed between "I" and "Name"-IRAP. However, only trial-type 1 of the name IRAP correlates with levels of the participant's self-esteem, that is, the IRAP effect of trial-type 1 "own name and positive words" in Name-IRAP changes depending on the level of self-esteem. Zhang and Tani. (2022) discussed that there was a functional difference between "own name" and "*watashi*" in that the evoking function (appetitive or aversive) of the name in making self-statement (e.g., I am good) was related to self-esteem, but "*watashi*" was not. They considered it reasonable that compared to "*watashi,"* "own name" is often used in relatively private or intimate conversations, which are more likely to contain emotional content. The findings indicated the lower the self-esteem, the weaker the trial-type 1 IRAP effect measured by the Name-IRAP.

The present study proposes to clarify the functional differences of "own name" depending on the level of self-esteem. From the findings of previous IRAP studies, it is assumed that STTDE appears only in the group with high self-esteem and is not observed in the group with low self-esteem for Name-IRAP. First, we administer the Japanese version of the Rosenburg Self-Esteem Scale (RSES-J; Sakurai, 2000) to groups with low and high self-esteem. Next, we employ the Name-IRAP setting that was used in the previous study (Zhang and Tani, 2022). Finally, we compare the IRAP data of higher and lower self-esteem groups and discuss the hypotheses that STTDE appears only in the group with high self-esteem for participant's name-positive trials and is not observed in the group with low self-esteem.

2. Methodology

(1) Participants and Setting

We used a flyer to recruit study participants from a local university; all were native Japanese speakers. The participants, 29 males and 65 females, ranged in age from 18 to 21 years. They were told the purpose of the research at the beginning of the study, and all signed informed consent documents; participants received ¥1,000 Amazon gift coupons as a reward. Participants completed the experiment individually in a quiet classroom with one of the researchers always present. This study

was approved by Ritsumeikan University’s Ethics Review Committee on Research with Human Subjects. According to the exclusion criteria described below, the end result was 16 participants (10 women and 6 men, age 18-21) in the group with higher self-esteem and 15 (6 women and 9 men, age 18-21) in the group with lower self-esteem.

(2) Apparatus and Materials

The experiment entailed administering the self-report RSES-J and Name-IRAP. The RSES-J (Sakurai, 2000) is a 10-item measure of self-esteem with each item rated on a 4-point Likert scale from 1 (strongly disagree) to 4 (strongly agree). The maximum possible score on the RSES-J is 40 points and the minimum is 10. Higher scores in RSES-J indicate higher self-esteem, and lower scores indicate lower self-esteem. The questionnaire demonstrated good internal validity and reliability (Sakurai, 2000).

Table 1. IRAP setting in this study. Created by the author

Consistent rule	
I am positive, others are negative.	
Inconsistent rule	
I am negative, others are positive.	
Label 1： Self	Label 1： Others
Participant′s name	Other′s name
Target 1“selfless words”	Target 2“selfish words”
English	English
good	bad
success	dishonest
honest	worthless
capable	nasty
pleasant	ashamed
confident	failure
Response Option 1	Response Option 2
Yes	No

Name-IRAP (Table 1) required participants to respond to various statements from the perspective of “I” or “others” where own name was label 1 and others was label 2. The target stimuli were positive (good, success, honest, capable, pleasant, and confident) and negative (bad, failure, dishonest, worthless, nasty, and ashamed) words, and the response options were confirming (*hai* in Japanese) and rejecting (*iie*). All stimuli were in Japanese for the following four trial types: 1. own name-positive, 2. own name-negative, 3. others-positive, and 4. others-negative (Figure 1).

(3) Procedure

First, all participants completed the self-report measure of the RSES-J. Participants were divided into both higher self-esteem and lower self-esteem groups based on their self-esteem scores. The criteria for grouping were the population mean (27.93) and standard deviation (5.51) calculated by Sakurai (2000), with participants who scored above “the mean + standard deviation (34.44)” as the higher self-esteem group and those who scored below “the mean - standard deviation (22.42)” as the lower self-esteem group, all other participants who scored from 23 to 34 were excluded from the analysis. In doing so, there were 21 in the group with higher self-esteem and 17 in the group with

lower self-esteem. Next, participants completed the Name-IRAP. The Name-IRAP provided four practice and three test blocks, with one consistent and one inconsistent trial in each block. The correct option was the exact opposite in the consistent and inconsistent trial. For example, in the participant's name-positive trial type, the participants need to select confirming on the consistent trial and rejecting on the inconsistent trial. The label stimulus appears at the top of the screen, the target stimulus appears in the middle, and the response options appear at the bottom left and bottom right of the screen. Participants pressed "D" to select the option on the left and "K" to select the option on the right. The positions of "confirm" and "reject" were randomized to ensure that they did not appear more than three times in a row in the same location.

To enter the next trial, the participant was required to select the correct response in a block of trials. A wrong response in a block of trials caused a red X to appear below the target stimulus. We conducted trial presentations with corrective feedback until each participant completed all 24 trial areas. In these 24 trials, each labeled stimulus was presented twice along with each target stimulus, and all trials were presented in random order.

There were four practices and three test blocks. Each block contained a consistent trial and an inconsistent trial. The correct option in the blocks of consistent trials was participant's name-positive/yes; participant's name-negative/no; others-positive/no; and others-negative/yes. In the blocks of inconsistent trials, the options were participant's name-positive/no; participant's name-negative/yes; others-positive/yes; others-negative/no. All participants were first presented with a block of consistent trials. After a respondent completed each trial, the IRAP gave the feedback on the correct rate and the average response time. At the end of the last block, the screen automatically switched to the text "Thank you for your participation."

Regarding mastery criteria, consistent with previous IRAP studies (e.g., Remue et al., 2013), we calculated the accuracy and average latency of each test block calculation. We excluded from the analysis data from participants who failed to achieve at least 80% accuracy or an average latency of less than 2500 ms in any test block. In addition to this, the reason for limiting the response latency is that if participants spend more time thinking, they will be able to respond after thinking properly, and the difference in response latency will be less significant. Also, during the thinking time, various noises (e.g., social desirability) are introduced into the participants' responses. On purpose, participants can react faster to socially desirable choices and slower to those that are not socially desirable. Also, the limitation on the percentage of correct responses implies a minimum level at which participants understand the rules and are willing to react correctly. With the above, the previous study suggests that if such a criterion appropriately met both requirements in the practice blocks, they immediately moved to the first pair of test blocks, but if participants failed to meet the requirements in all four practice blocks, then IRAP was terminated, and these data were discarded. In applying these exclusion criteria to all test blocks, we excluded responses from a total of 38 participants. Of the 31 participants we analyzed, there were 16 in the group with higher self-esteem and 15 were in the group with lower self-esteem. We explained the IRAP test to each participant to ensure that they understood the testing process prior to conducting the test.

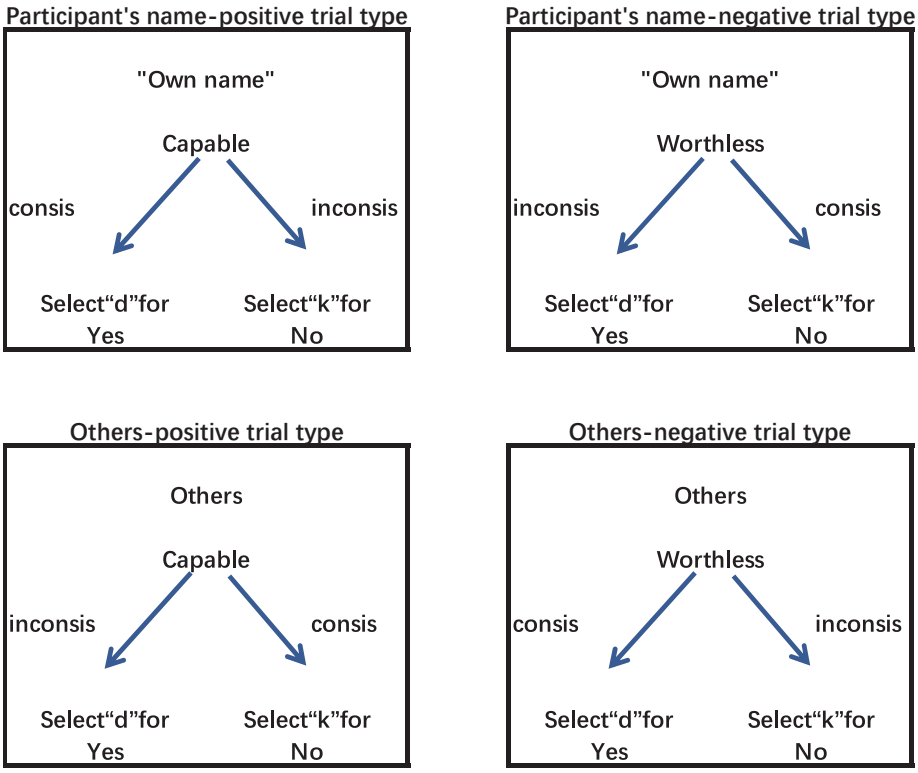

Figure 1. The four trial-types used in the two IRAPs. Created by the author

Note. Consis = correct response in the consistent block; inconsis: correct response in the inconsistent block.

3. Results

(1) Equality Analysis of Grouping

To analyze the equality of the groupings, we analyzed the difference in the age and sex of the two groups. t-test analysis showed no significant age differences ($t = -.92$, $df = 29$, $p = .37$, *n.s*). The chi-square test showed no gender-significant difference for the two groups($\chi 2(1) = 1.57$, $p = .21$, *n.s*). Based on these results, the two groups were found to be equitable in age and gender.

(2) DIRAP Scores

The DIRAP score is the difference between the mean response times for the first correct selections in each block divided by the standard deviation of all the response times in both blocks (see Hussey et al., 2015). We calculated DIRAP scores for each of the four trial types (see Barnes-Holmes et al., 2010a), and Figure 2 shows the scores for the IRAP with the two groups with whom we conducted this study. Positive DIRAP scores indicated responses consistent with the rule that I am positive, or others are negative, reflected as choosing “yes” more quickly than “no” in the self-positive and others-negative trials, and choosing “no” more quickly than “yes” in the self-negative and others-positive trials in both the consistent and inconsistent blocks. Negative DIRAP scores indicate responses consistent with the rule that I am negative, or others are positive, reflected as choosing “no” more quickly than “yes” in the self-positive and others-negative trials and choosing “yes” more quickly than “no” in the self-negative and others-positive trials in both the consistent and inconsistent blocks.

(3) IRAP for the Higher Self-esteem Group

In Figure 2, the higher self-esteem group shows results for the higher self-esteem group. Participants showed 0.51 DIRAP scores in the participant's name-positive, 0.12 DIRAP scores in the participant's name-negative, −0.01 DIRAP scores in the others-positive, and −0.18 DIRAP scores in the others-negative trials. That is, participants selected "yes" more quickly than "no" in the participant's name-positive and others-positive trials, "no" more quickly than "yes" in the participant's name-negative and others-negative trials.

We used a one-sample t-test to analyze whether the DIRAP scores in IRAP were significantly different from zero and found a significant difference for the participant's name-positive trial type (t = 6.14, df = 15, $p < .001$, $\eta\rho 2 = .33$). Participant's name-negative ($t = .93$, df = 15, $p = .37$, $\eta\rho 2 = .51$), others-positive ($t = -.15$, df = 15, $p = .89$, $\eta\rho 2 = .34$), and others-negative ($t = -1.70$, df = 15, $p = .11$, $\eta\rho 2 = .43$) did not reach significance.

We used one-way repeated-measures ANOVAs to investigate the differences in the IRAP effect for all trial types and identified a significant main effect of trial type (F (3, 15) = 10.35, $p < .001$, $\eta\rho 2 = .41$). Bonferroni post hoc tests indicated that the effect for participant's name-positive trials differed significantly from those for the participant's name-negative ($p = .012$), the others-positive ($p = .009$), and the others-negative trials ($p < .001$). The remaining trial types did not differ significantly from each other ($p > .05$). Thus, STTDE emerged in name-positive trials for higher self-esteem group.

(4) IRAP for the Lower Self-esteem Group

In Figure 2, the lower self-esteem group shows results for the lower self-esteem group. Participants showed 0.01 DIRAP scores in the participant's name-positive, −0.04 DIRAP scores in the participant's name-negative, −0.17 DIRAP scores in the others-positive, and −0.18 DIRAP scores in the others-negative trials. That is, participants selected "yes" more quickly than "no" in the participant's name-positive, the participant's name-negative and others-positive trials, and "no" more quickly than "yes" in others-negative trials.

One-sample t-test confirmed all trial types did not reach significance. The participant's name-positive trial type (t = 15, df = 14, $p = .88$, $\eta\rho 2 = .37$), participant's name-negative ($t = -.45$, df = 15, $p = .66$, $\eta\rho 2 = .34$), others-positive ($t = -2.49$, df = 15, $p = .16$, $\eta\rho 2 = .27$), and others-negative ($t = -2.18$, df = 15, $p = .27$, $\eta\rho 2 = .31$).

One-way repeated-measures ANOVAs showed no significant main effect of trial type (F (3, 14) = 1.30, $p = .29$). Bonferroni post hoc tests indicated that the effect for all trials did not differ significantly from each other ($p > .05$). Thus, no STTDE was found for lower self-esteem group.

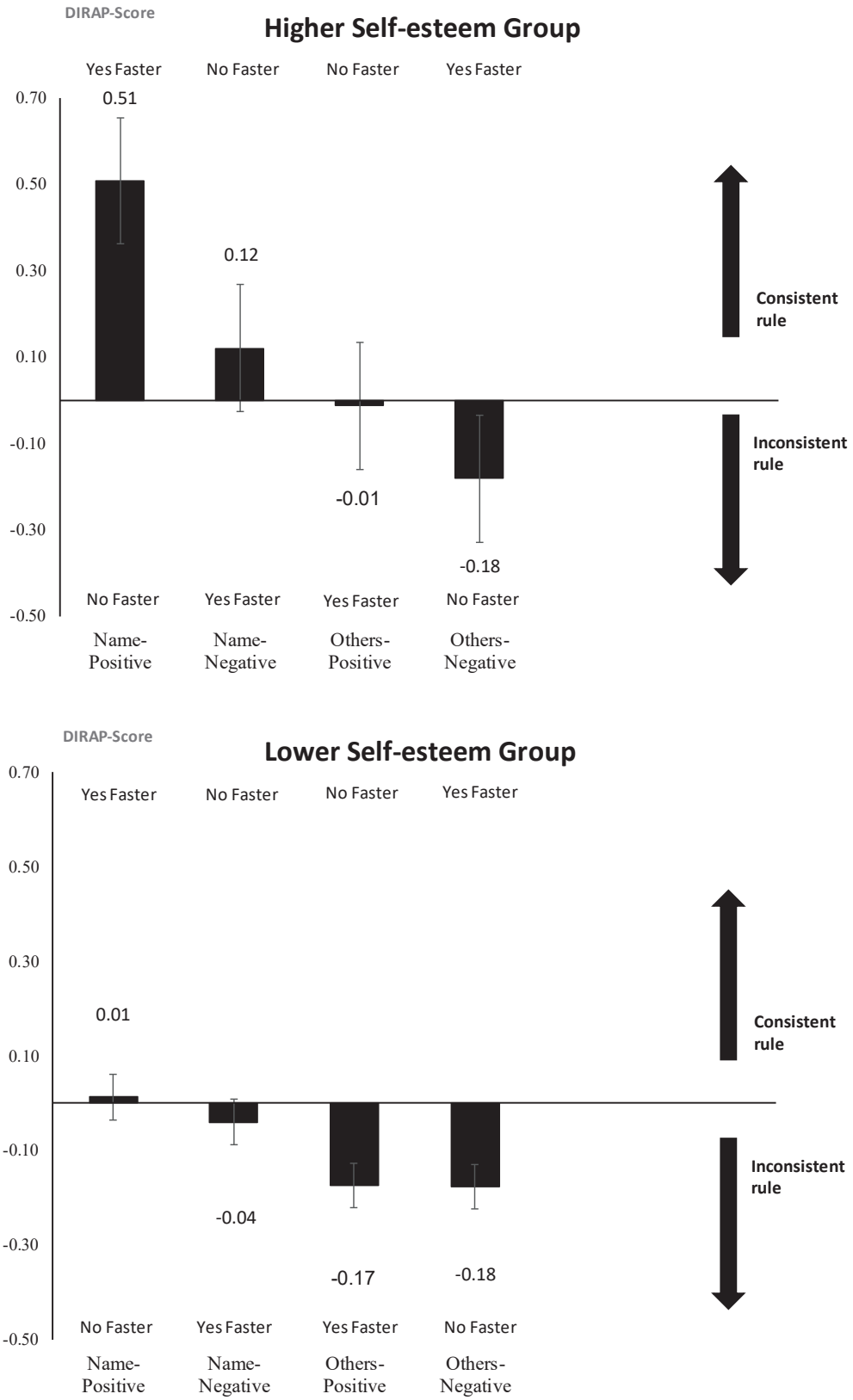

Figure 2. Mean DIRAP scores for the four trial-types on name-IRAP. Created by the author
Note. DIRAP scores for trial-types 3 and 4 in both IRAPs are inverted to create a common axis across all trial-types.

4. Discussion

This study investigated the functional differences of "own name" depending on the level of self-esteem in IRAP by observing the results of one-sample t-tests and ANOVAs of two higher and lower self-esteem groups. First, regarding the one-sample t-tests results of IRAP for two groups, only the DIRAP scores for the higher self-esteem group showed a significantly different response that participant's response "yes," was faster than "no" in participant's name-positive trials. Second, regarding the ANOVAs results of the two groups, only the IRAP effects for the higher self-esteem group identified a significant main effect of trial type, the IRAP effects for the participant's name-positive trials differed significantly from remaining trial types. However, no significant differences in IRAP effects were found for the lower self-esteem group. The result confirmed the hypothesis of the present study that STTDE appears only in the group with high self-esteem and is not observed in the group with low self-esteem. From these outcomes, we suggest that the function of "own name" is different depending on the level of self-esteem for the Japanese language.

(1) STTDE appears only in the group with high self-esteem

Let us consider the ANOVAs results of the IRAPs. An important finding from this study that we wish to emphasize is that according to our hypothesis, STTDE (the size of the IRAP effect for one trial type appears to dominate over the other three trial types) appears only in the group with high self-esteem for participant's name-positive trials and is not observed in the group with low self-esteem. The results of ANOVAs confirmed the hypothesis. At this point, it is important to emphasize that STTDE has been observed several times in other IRAP studies, and its appearance has been explained by the fact that all the stimuli used in IRAP reach a certain specific condition. That is, although both groups used the same stimulus set in the IRAP, the higher self-esteem group achieved this particular condition, while the lower self-esteem group did not. By analyzing this result through DAARRE, it seems to be clearer what kind of changes in participant's name related to the levels of self-esteem that were produced.

(2) Analysis of the DAARRE model

As mentioned in the Introduction, the DAARRE model explains the IRAP effect related to the interactions between the functional properties of the individual stimuli (e.g., the extent to which they evoke appetitive or aversive) and the relationships between the label, target stimuli and the functional properties of the response options (e.g., the relationship between the two stimuli are consistent or inconsistent with their learning history (Barnes-Holmes et al., 2020). According to the explanation of STTDE, Finn et al. (2018) reported on a color-shape IRAP in which the labels were colored and shaped words, and the targets were words denoting colors and shapes, they identified the STTDE in the significantly higher DIRAP scores for the trial-type that presented the label color, and the target color words. The authors attributed their findings to the stronger orienting functions of color words due to their higher frequency of occurrence in the language than shape words. Consistent with this explanation, the authors employed the words, "spoon" and "fork," with approximately equal frequencies in the language (that is, with equal orienting functions), and the targets were pictures of spoons and forks (equal orienting functions). The results showed that the STTDE did not occur in the spoon-fork IRAP (Finn et al., 2018). Based on these outcomes, Barnes-Holmes et al. (2020) explained that STTDE only emerged when participants experience a type of "yes-yes" effect in DAARRE model

analysis. That is, in the stronger orienting or evoking function related to the “yes” response, the relation between two stimuli is consistent with participant’s learning history related to the “yes” response.

The DAARRE model can explain the results of the current study for IRAPs. According to the analysis of the DAARRE model, the positive (+) and negative (−) labels indicate the relative strength of the orienting or evoking function for label and target stimuli, and the relation between two stimuli is consistent (+) or inconsistent (−) with the participant’s learning history. It is critical to note that Barnes-Holmes et al. (2020) suggested that a stronger orienting or appetitive evoking function (marked as Cfunc +) related to the confirming or approach behavior (“yes” response). Meanwhile, a weaker orienting or aversive evoking function (marked Cfunc as −) related to a “no” response. The relation is consistent with the participant’s learning history (marked as Crel +) related to a “yes” response, and inconsistent (marked as Crel −) related to a “no” response.

In relation to the present study, as shown in Figure 3, first, individuals tended to make more self than other statements, and label 1 “own name” appeared to have a stronger orienting function (marked as +) than that of label 2, “others” (marked as −). Regarding the evoking function, Pidgeon et al. (2021) found that pictures of failure evoked a strong aversive function regarding the self among participants with low self-esteem. Label 1 “own name” stimuli likely possesses an appetitive evoking function (marked as + next to the stimuli) for participants with higher self-esteem, and an aversive evoking function (marked as −) for participants with lower self-esteem. Label 2 “others” may possess a neutral evoking function (marked as ±). Second, previous studies (Leech et al., 2016; Leech et al., 2017; Barnes-Holmes et al., 2020; Pidgeon et al., 2020; Barnes-Holmes et al., 2021; Barnes-Holmes et al., 2022) suggested that “positive word” likely possesses positive Cfunc for orienting and evoking properties (marked as +), in contrast with those of negative words (marked as −). The relationship between the label and target stimuli (Crel) indicated whether a participant’s IRAP results were consistent with that participant’s learning history. Third, the expectation is that individuals will more strongly associate “self” with positive rather than negative words for participants with higher self-esteem. So, the participant’s name-positive relationships were indicated with positive (marked as +, consistent with learning history) and the participant’s name-negative relationships with negative (marked as −, inconsistent with learning history) for participants with higher self-esteem. On the other hand, for the participant’s name-positive relationships with positive (marked as −) and the participant’s name-negative relationships with negative (marked as +) for participants with lower self-esteem, we predict that there will not be great differences between Crel for others-positive and others-negative relationships in the participant’s learning history so that Crel could possess the same strength and be represented as both ±. Finally, “yes” is typically used in natural language to indicate coherence, and “no” indicates incoherence.

Typically, in works by previous IRAP researchers (Finn et al., 2018; Barnes-Holmes et al., 2020; Kavanagh et al., 2019), the STTDE emerged in the trial types that produced positive functions for all Crels and Cfuncs (maximum coherence). Furthermore, Barnes-Holmes et al. (2020) indicate the stronger evoking functions (appetitive or aversive) that are likely to dominate for individuals who are more afraid of spiders. Related to the current study, the stronger evoking function is likely to dominate for individuals with low as opposed to high self-esteem. That is, participants’ responses will be dominated by the evoking function rather than by the orienting function if the stimuli have a stronger appetitive (aversive) function. As mentioned, participants may experience a type of “yes-yes” effect when presented with this maximum coherence trial-type. Based on these findings, in the analysis of

the DAARRE model sample for two groups, a response of "yes" in the participant's name-positive trials indicates a type of "yes-yes" effect (in evoking dominating, two of evoking Cfunc and Crel all marked as +). In contrast, participants may experience a type of "yes-no" or "no-no" effect for the remaining trial-types (reduced coherence).

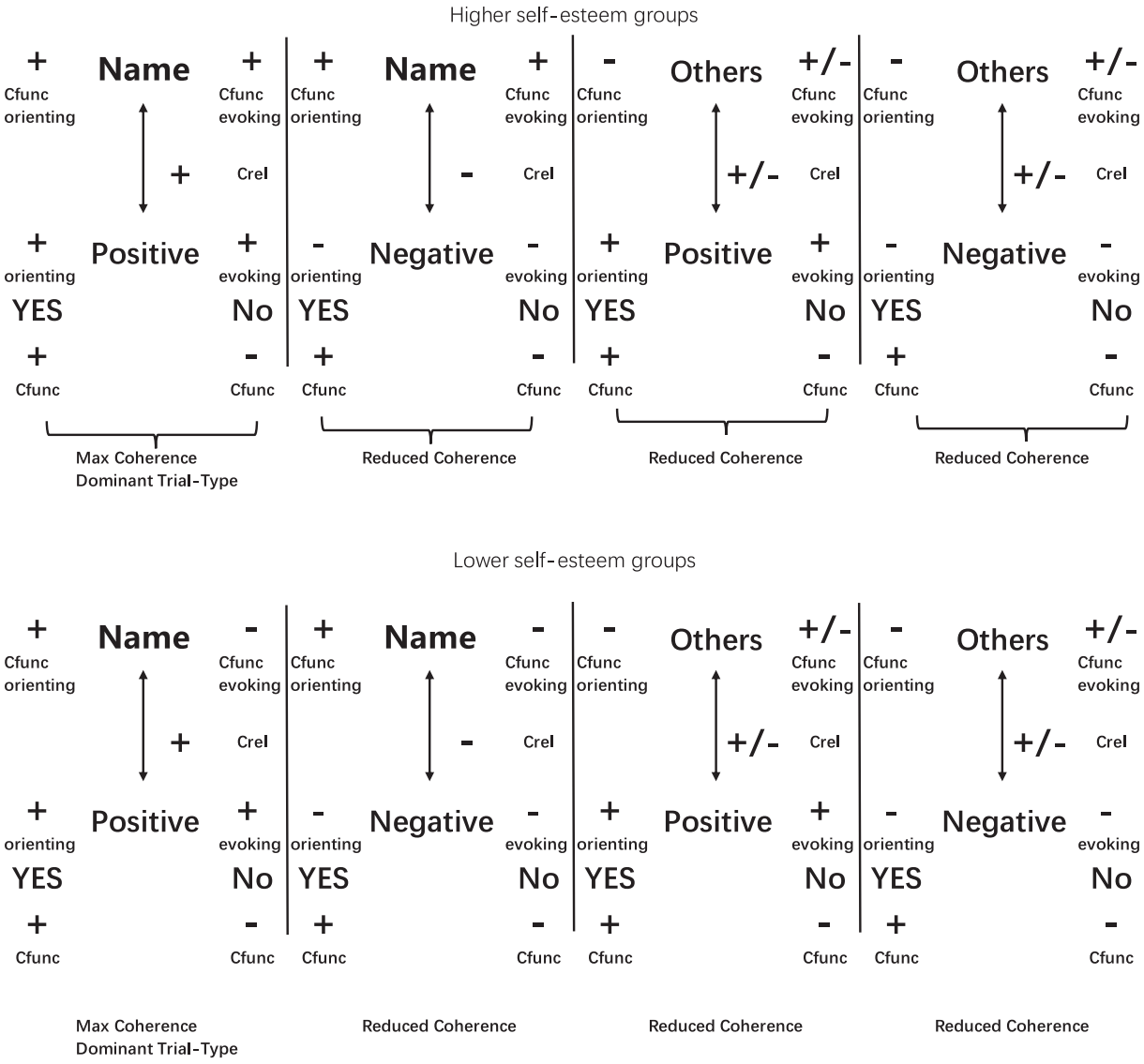

Figure 3. The DAARRE model as it applies to the IRAP. Created by the author

Note. The positive (+) and negative (−) labels refer to the relative positivity of the Cfuncs for each label and target, the relative positivity of the Crels, and the relative positivity of the RCIs in the context of the other Cfuncs, Crels, and RCIs in that stimulus set.

Furthermore, regarding the one-sample t-test results of the IRAPs, there is a question that if the function of "own name" differed between higher and lower self-esteem participants, why did only one of the participant's name-positive trials perform differently? And why did the participant's name-negative trial not perform differently? The explanation of the DAARRE model analysis would clarify this question. Figure 4 shows the DAARRE model analysis findings for individuals with high and low self-esteem in the participant's name-positive and participant's name-negative trial-types in the name-IRAP. For the participant's name-positive trial-type, the maximum coherent response for participants with higher self-esteem produces a maximum coherence ("yes-yes" effect) IRAP effect that related to

higher DIRAP scores. However, participants with lower self-esteem may experience a type of "no-no" effect related to smaller or weaker IRAP effects. Thus, the function of "own name" differed between higher and lower self-esteem participants leading to the different DIRAP scores. In the participant's name-negative trials, although the evoking function of "own name" changed depending on the self-esteem, higher and lower self-esteem produced the relatively similar "yes-no" effect which related to the same DIRAP scores. Therefore, only one of the participant's name-positive trials performed differently in DIRAP scores.

Overall, with regard to the hypothesis of this present study that is based on the previous study (Zhang and Tani, 2021) suggestion, "STTDE appears only in the group with high self-esteem and is not observed in the group with low self-esteem," these outcomes supported the hypothesis. The results suggest that the evoking function (appetitive or aversive) of the name was related to the level of self-esteem. It is important to emphasize that this is the second study to successfully control the appearance and disappearance of STTDE by manipulating the variables in a number of IRAP studies. Furthermore, it is the first study to successfully control STTDE by changing other variables than IRAP stimuli settings, which is of great significance for the development of IRAP studies.

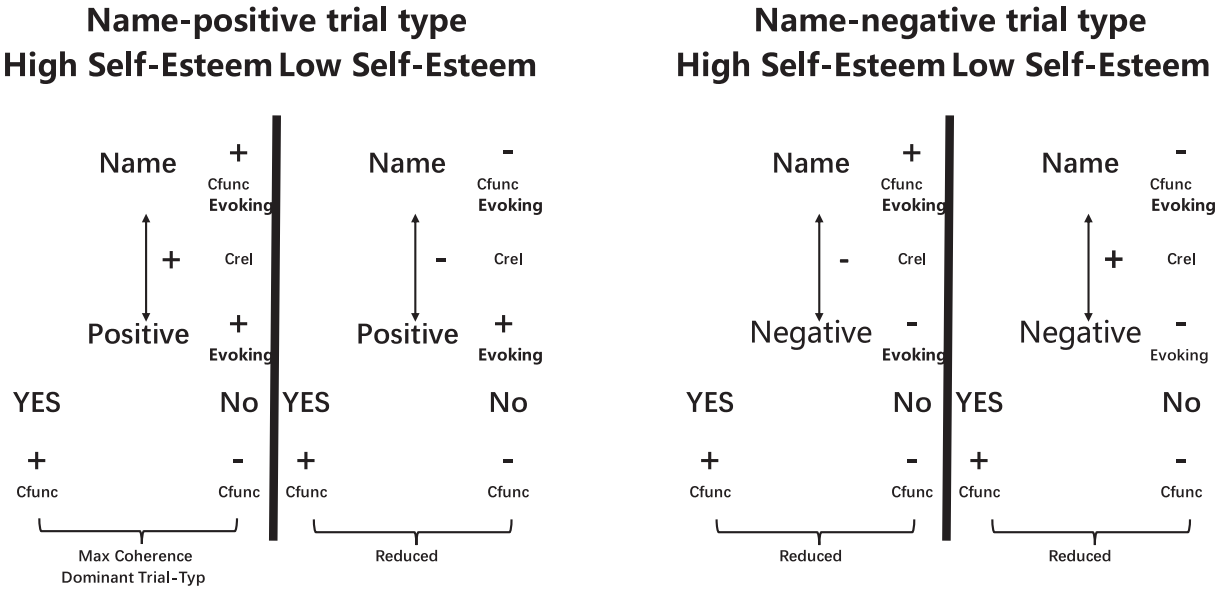

Figure 4. The DAARRE model as it applies to individuals with high and low self-esteem. Created by the author
Note. For the participants with low self-esteem, + indicates an aversive evoking function of their names. For those with high self-esteem, − indicates that their names evoked an appetitive function. For Crel, + indicates consistency with a participant's learning history, and − indicates inconsistency.

(3) Limitations

To close this study, we wish to consider its limitations. First, although the current study successfully identified differences in the functioning of participants' name and positive (negative) words, we cannot determine whether this change in function is produced only when "own name" and the evaluative words (positive and negative) are used together, or whether "own name" alone will produce it. Second, we used full name to identify "self" in the Japanese language. However, some participants reported that it would be easier to respond if using only one's first name (e.g., Taro and positive words). It may be necessary to reconduct this study with a different "own name" selection. Third, we noticed that there is too much difference between the initial male/female ratio and the ratio

of male/female subjects in the experiment. The current study was grouped by "Japanese adolescents mean ± standard deviation" of self-esteem scores. In other words, all scores close to the mean of self-esteem scores were excluded. The results of the grouping suggest that females tended to score closer to the mean than males among this study's participants. In future studies, gender differences in self-esteem scales may need to be taken into account. Finally, one unique feature of Japanese culture is that the self-meaning words are different depending on gender (*watashi*, *ore*, and *boku*, for instance). Gender difference may be a very important variable in the Japanese self-concept, and so the literature would benefit from future research on these gender differences.

4. Conclusion

This study investigated the functional differences of "own name" depending on the level of self-esteem in IRAP. The hypothesis that STTDE appears only in the group with high self-esteem and is not observed in the group with low self-esteem is supported by the results. This study successfully identified the difference that higher self-esteem was related to a stronger appetitive evoking function, and lower self-esteem to a stronger aversive evoking function for own name in Name-IRAP. It is critical to note that the current study identified the evoking function of own name changes in an "I am/am not positive/negative" verbal frame. The study also successfully controls the Appearances of STTDE as a first IRAP study for Japan. As mentioned in the Introduction, acceptance and commitment therapy notes not that "I am not bad," but that "I had a thought that I am bad," is important for a sense of self because it empowers well-being (Luciano et al., 2011; Kavanagh et al., 2020). Based on the results of the current study, it suggests not using "…important for me (*watashi*)" but "…important for Taro (own name)" may be more effective for the Japanese to notice what is appetitive or aversive for him/her in future psychological clinical works.

References

Barnes-Holmes, D., Barnes-Holmes, Y. and McEnteggart, C. 2020. Updating RFT (more Field than Frame) and its Implications for Process-based Therapy. *The Psychological Record*, 70(4), 605-624. https://doi.org/10.1007/s40732-019-00372-3

Barnes-Holmes, D., Barnes-Holmes, Y., McEnteggart, C., and Harte, C. 2021. Back to the Future with an Updated Version of RFT: More Field than Frame? *Perspectivas em Análise do Comportamento*, 12(1). https://doi.org/10.18761/PAC.2021.v12.RFT.03

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R. and Stewart, I. 2006. Do You Know What You Really Believe? Developing the Implicit Relational Evaluation Procedure (IRAP) as a Direct Measure of Implicit Beliefs. *Irish Psychologist*, 32(7), 169-177.

Barnes-Holmes, D., Barnes-Holmes, Y., Luciano, C., and McEnteggart, C. 2017. From IRAP and REC Model to a Multi-dimensional Multilevel Framework for Analyzing the Dynamics of Arbitrarily Applicable Relational Responding. *Journal of Contextual Behavioral Science*, 6(4), 473-483. https://doi.org/10.1016/j.jcbs.2017.08.001.

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., and Boles, S. 2010a. A Sketch of the Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and Coherence (REC) Model. *Psychological Record*, 60(3), 527-542. https://doi.org/10.1007/BF03395726

Barnes-Holmes, D., Murtagh, L., Barnes-Holmes, Y., and Stewart, I. 2010b. Using the Implicit Association Test and the Implicit Relational Assessment Procedure to Measure Attitudes Toward Meat and Vegetables in Vegetarians and Meat Eaters. *Psychological Record*, 60(2), 287-306. https://doi.org/10.1007/BF03395708

Barnes-Holmes, D. and Harte, C. 2022. Relational Frame Theory 20 Years on: The Odysseus Voyage and Beyond. *Journal of the Experimental Analysis of Behavior*, 117(2), 240-266.

Finn, M., Barnes-Holmes, D. and McEnteggart, C. 2018 Exploring the Single-trial-type-dominance-effect on the IRAP: Developing a Differential Arbitrarily Applicable Relational Responding Effects (DAARRE) Model. *The Psychological Record*, 68, 11-25.

Hayes, S. C. 2004. Acceptance and Commitment Therapy, Relational Frame Theory, and the Third Wave of Behavior Therapy. *Behavior Therapy*, 35, 639-665.

Hirose, Y., and Hasegawa, Y. 2016. *Japanese People from the Perspective of the Japanese Language* Tokyo: Kaitakusha. (In Japanese)

Hughes, S., Barnes-Holmes, D., and De Houwer, J. 2011. The Dominance of Associative Theorizing in Implicit Attitude Research: Propositional and Behavioral Alternatives. *Psychological Record*, 61(3), 465-496. https://doi.org/10.1007/BF03395772

Hussey, I., Thompson, M., McEnteggart, C., Barnes-Holmes, D., and 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

Kavanagh, D., Barnes-Holmes, Y., and Barnes-Holmes, D. 2020. The Study of Perspective-taking: Contributions from Mainstream Psychology and Behavior Analysis. *Psychological Record*, 70(4), 581-604. https://doi.org/10.1007/s40732-019-00356-3

Kavanagh, D., Matthyssen, N., Barnes-Holmes, Y., Barnes-Holmes, D., McEnteggart, C., and Vastano, R. 2019. Exploring the Use of Pictures of Self and Other in the IRAP: Reflecting upon the Emergence of Differential Trial-type Effects. *International Journal of Psychology & Psychological Therapy*, 19(3), 323-336.

Leech, A., Barnes-Holmes, D., and Madden, L. 2016. The Implicit Relational Assessment Procedure (IRAP) as a Measure of Spider Fear, Avoidance, and Approach. *The Psychological Record*, 66(3), 337-349.

Leech, A., Barnes-Holmes, D., and McEnteggart, C. 2017. Spider Fear and Avoidance: A Preliminary Study of the Impact of Two Verbal Rehearsal Tasks on a Behavior–behavior Relation and its Implications for an Experimental Analysis of Defusion. *The Psychological Record*, 67(3), 387-398.

Luciano, C., Ruiz, F. J., Vizcaíno Torres, R. M., Sánchez Martín, V., Gutiérrez Martínez, O. and López-López, J. C. 2011 A Relational Frame Analysis of Defusion in Acceptance and Commitment Therapy. A Preliminary and Quasi-Experimental Study with At-Risk Adolescents. *International Journal of Psychology and Psychological Therapy*, 11, 165-182.

Feltham, C. and Horton, I. 2006 *The Sage Handbook of Psychotherapy and Counselling* (2nd ed.) (pp. 292-297). London, UK: Sage.

Pidgeon, A., McEnteggart, C., Harte, C., Barnes-Holmes, D. and Barnes-Holmes, Y. 2021. Four Self- Related IRAPs: Analyzing and interpreting effects in light of the DAARRE model. *Psychological Record*, 71(3), 397-409. https://doi.org/10.1007/s40732-020-00428-9

Ruiz, F. J. 2010. A Review of Acceptance and Commitment Therapy Empirical Evidence: Correlational, Experimental Psychopathology, Component and Outcome Studies. *International Journal of Psychology and Psychological Therapy*, 10, 125-162.

Remue, J., Houwer, J. D., Barnes-Holmes, D., Vanderhasselt, M., and Raedt, R. D. 2014. To Be or Want to Be: The Role of Actual Versus Ideal Self in Implicit Self-esteem. *PLOS ONE*, 9. https://doi.org/10.1371/journal.pone.0108837

Sakurai, S. 2000. Investigation of the Japanese Version of Rosenberg’s Self-Esteem Scale. *Developmental and Clinical Psychology*, 12, 65-71.

Zhang, P. and Tani, S. 2022. Examination of the Function of Self-meaning Stimulus Used for IRAP in Japanese. Poster Presented at *Association for Contextual Behavioral Science 2022*, 2022/6/14-19, San Francisco.