The Effect of Defusion on the Believability of Japanese Speakers' Self-Descriptions: Results of the Function Acquisition Speed Test

Maho KONDA* and Shinji TANI**

Abstract: This study investigates how defusion affects the believability of Japanese speakers' negative and positive self-statements and uses the Function Acquisition Speed Test (FAST) to examine the detailed process that underlies such an effect from a relational frame perspective. Twenty-three participants were randomly assigned to either the defusion or the non-defusion group. First, the participants reported negative and positive self-statements, but only the defusion group completed a defusion exercise and practiced it for one week. After the week, all participants took the FAST and rated the believability of the reported self-statements. Results showed that there were no significant differences in the FAST effect, the mean believability of negative and positive self-statements between the defusion and non-defusion groups, although the mean believability of negative self-statements was positively correlated with positive self-statements only in the non-defusion group, suggesting an insufficient defusion effect as well as failure to identify the defusion remained unclear. The implications, limitations, and potential future directions were discussed.

Keywords: *Function Acquisition Speed Test, relational frame theory, defusion, self-statements*

1. Introduction

(1) Human Suffering and Acceptance and Commitment Therapy

Psychotherapists rely on the power of language when they intervene in their client's life. Psychotherapists speak articulately, try to understand the client's story, and orient the client to what they should notice in order to provoke insights and so on. Language itself works as an intervention, so understanding the language process is critical for psychotherapists (Villatte et al. 2016). Language is a unique ability of human beings. According to the Relational Frame Theory, it allows humans to relate and respond to a stimulus symbolically through language in terms of their learning history (Villate et al. 2016). They can derive the relation between stimuli A and C despite only having the knowledge that stimulus A equals stimulus B, which equals stimulus C, and despite the absence of any physical stimuli. Similarly, people can symbolically presume the equivalent relation

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

** Professor, College of Comprehensive Psychology, Ritsumeikan University

Email: * gr0419vi@ed.ritsumei.ac.jp

** stt23197@pl.ritsumei.ac.jp

Received on August 28, 2023, accepted after peer reviews on April 4, 2024; published online: June 25, 2024.

Asia-Japan Research Academic Bulletin, 2024

ONLINE ISSN 2435-306X, Vol.5, 11

of "I" and "worthless" from only learning that they do not have a lot of money and that a person with no money is worthless. This symbolic way of relating and responding to stimuli based on one's learning history is called arbitrarily applicable relational responding (AARR), which occasionally causes problematic behaviors by changing a stimulus function in a maladaptive way, which can inhibit one from acting toward the enrichment of their well-being or values (McHugh, Stewart, and Priscilla 2019). Particularly, inflexible self-related AARR narrows a variety of behavioral repertoires that drastically reduce the likelihood of engaging in adaptive behaviors in a context (Villatte et al. 2016). For instance, in a situation where a person believes they are worthless because they do not have a lot of money (I = do not have money = worthless), the function of "worthless" transforms that of "not having a lot of money" and that of "I." The latter two used to perform neutral functions, but relating them to "worthless" can transform them in such a way that they evoke negative feelings. The person may feel that "I" (they) cannot do anything such as care for their family, enjoy their life, and have new friends because of the transformation of the stimulus function of "I" given by AARR with "worthless." As a result, this person may give up engaging in family activities, devoting time to hobbies, or going outside to meet new people, which contributes to the self-concept becoming rigid and being treated as literal truth (Atkins and Styles 2015). This change due to AARR is called the transformation of a stimulus function and is considered to be among the factors triggering overcontrolling behaviors in a maladaptive way that causes human suffering (Villatte et al. 2016). Simply put, people interpret everything based on what they have already learned, and sometimes, these interpretations heavily control what they want to do, inhibiting them from living their desired lives. This perspective has contributed to the development of a form of psychotherapy called Acceptance and Commitment Therapy (ACT).

ACT, the third-wave cognitive behavioral therapy that focuses on AARR, is an empirically supported treatment for chronic pain, depression, mixed anxiety disorders, psychosis and obsessive-compulsive disorder (Society of Clinical Psychology 2022). ACT does not target the problematic behaviors, emotions, feelings, or thoughts themselves due to AARR, but rather, it attempts to slow down or disrupt the transformation of a stimulus function by analyzing the context from which these issues arise (i.e., altering the relation oneself = worthless without changing its content by simply adding the prefix of "I am having a thought" in front of the relation).

(2) Defusion

Defusion, a concept under ACT, is frequently implemented to address human suffering and is known to be effective in mitigating the believability of negative self-statements, the willingness to experience negative self-statements (Masuda et al. 2004), smoking behaviors (Beadman et al. 2015), and learned helplessness (Hooper and McHugh 2013). However, scholars have yet to explain how defusion can yield certain effects, which may be due to the confusion surrounding its definition. To illustrate, the concept of defusion can be defined from three aspects: procedure, outcome, and process (Assaz et al. 2018). Defusion refers to the manipulation of a context, such as labeling your thoughts (LYT) and word repeating techniques (WRT), while defusion outcome entails changes in an individual's traits or behavioral tendencies. Studies classified under defusion outcome include Masuda et al. (2004), who implemented a WRT in which individuals repeated the word in their negative thought that had the most detrimental impact on them for 30 seconds (e.g., repeating "fat" in the thought "I am fat") and found that the method was more effective in reducing emotional discomfort and the believability of the thought than attempts to distract oneself or control their negative emotion. Healy et al. (2008) investigated the effect of LYT on negative and positive self-statements. The defused negative self-statements "I am having a thought that XXX" (e.g., "I am

having a thought that I am a bad person") were rated lower in emotional discomfort and higher in willingness and believability of the statements compared to the normal presentation of the statement (e.g., "I am a bad person") or abnormal presentation (e.g., "I have a wooden chair and I am a bad person"). However, a limited effect of defusion on emotional discomfort, willingness, and believability was found in the context of positive self-statements. The effect on believability contradicts Masuda et al. (2004) because the participants did not rate the self-statements themselves; they rated, for instance, the whole statement of "I am having a thought that I am a bad person." In effect, the believability of defused self-statements rated higher than the others. Otherwise, the finding of the believability of self-statements was expected to be consistent with the study of Masuda et al. (2004), yet this has not been tested.

Defusion is known to mitigate learned helplessness and tobacco use as well as increase behavioral fluency. Meanwhile, the defusion process refers to changes in the interaction between an individual and their environment. Blackledge (2007) argued that the defusion process pertains to focusing on a context and disrupting one's verbal behaviors; hence, the stimulus function is reduced. Comparing the defusion process with defusion outcome, however, the defusion process has not been well studied because the defusion concept involves a defusion (procedure) that can yield a defusion (outcome) via a defusion (process), which makes it difficult to focus on and distinguish the process of defusion from other aspects.

Today, for understanding the process of changes, including the defusion process has been vital, especially for advanced ACT therapists, as it allows them to avoid using the defusion process for specific symptoms just because it is the primary or typical technique for treating them. Since each client experiences different processes, the feasibility of the defusion process depends on the interaction between their ongoing process and the defusion process. Therefore, therapists who implement defusion must understand the defusion process well.

(3) Methodology For Investigating the Defusion Process

Contextual behavioral science, in which ACT is rooted, highlights the need for more behavioral measures to identify the process of changes because this would enhance one's understanding of the basic behavioral accounts of the process of change that must link concepts to the context to be fully functional. Indeed, the defusion process was not fully investigated via behavioral measurement. The word "defusion" refers to various aspects of the concept of defusion. Different researchers have used "defusion" to describe the defusion process itself or its outcomes or the process of changes resulting from it. The definition of defusion is generally defined as attending to a context and disrupting one's verbal behaviors; therefore, the transformation of a stimulus function (process) results in, for instance, a reduction in the believability of negative self-statements as well as an increase in behavioral fluency and so on (outcomes) through context manipulation (procedure). Therefore, it is reasonable to implement a functional behavioral assessment tool that measures the effect of the transformation of a stimulus function on the resulting behaviors. Two methods have been developed in this regard: Implicit Relational Assessment Procedure (IRAP) and the Function Acquisition Speed Test (FAST).

IRAP, a computer-based assessment tool, takes advantage of humans' tendency to respond faster to stimulus relations in a way that is consistent with their learning history in the presence of time constraints rather than an inconsistent way with their knowledge acquisition history. Thus, IRAP requires participants to provide correct responses with regard to both their consistent and inconsistent relations with their learning history and compare their response latency. The former is called the consistent block, while the latter is called the inconsistent block. Typically, participants

are shown two stimuli (e.g., "I" and "success") with other stimuli that define their relation (e.g., "similar"), and they must respond to the stimuli by pressing keys (e.g., answering "yes," or "no"). They then either receive a "correct" feedback message if their response is correct or simply proceed to the next trial with no feedback in the case of an incorrect response. The difference in response latency between the consistent and inconsistent blocks will be analyzed, and when the response latency in a block is shorter, the stimulus relations tested in the block are consistent with one's learning history. While this is an extremely useful way to assess AARR, the transformation of the stimulus function, and the ensuing behaviors, IRAP has pointed out several issues for the use of response latency as the indicator of the strength or stability of any instance of behaviors, a post-hoc data cleansing method for acquiring data stability, imbalanced feedback, and the presentation format that lead participants to predict what was assessed in the procedure (Cummins et al. 2018; O'Reilly et al. 2012).

The FAST is a computer-based assessment tool (O'Reilly et al. 2012) that capitalizes on human tendency, demonstrating that stimulus relations functionally coherent to their learning history are easier to learn than those functionally incoherent to their knowledge acquisition history. Therefore, the FAST requires participants to learn stimulus relations that are both functionally coherent and functionally incoherent to their learning history. Participants who undergo the FAST only see one stimulus and respond by pressing either M or Z on the keyboard. For instance, the researcher, assuming participants regard themselves as more successful than others, asks them to press the M key in the presence of a stimulus "I" or "success" and press the Z key in the presence of a stimulus "other" or "fail." Feedback, either "Correct" or "Incorrect," will be given after the response, and a new trial begins. By repeating this task, participants established the response class of pressing the M key for "I" and "success" and pressing the Z key for "other" and "fail" based on given feedback. Easier learning indicates the response class learned here is consistent with the hypothesized participant's learning history. Therefore, the block in which this learning occurs is named the "consistent block." FAST also includes a block where the participant's response is reversed; that is, pressing the M key in the presence of "I" and "fail" and pressing Z in the presence of "other" and "success" are reinforced via feedback. This block is named the "inconsistent block" as participants learn reversed or inconsistent response class with their hypothesized learning history. Easier learning in the consistent block than the inconsistent block suggests a stronger learning history that is consistent with the response class established in the consistent block. The FAST index, Strength of Relation (SoR), was used to access participants' easiness of learning in the consistent block. The positive SoR indicates easier learning of the consistent block, and 0 or negative SoR indicates the reverse tendency.

FAST has four advantages over IRAP. First, FAST analyzes the response class acquisition rate (i.e., SoR), and the fluency-type measure, which is indicative of the strength or stability of behaviors. Second, FAST acquires data stability by improving experimental control over response variability via modification of the procedure rather than post-hoc process. Third, FAST gives participants feedback for correct and incorrect responses. It can get rid of ambiguity about whether the greater learning speed in one block is produced by the assumption of the FAST procedure or other reasons such as something that was raised in IRAP's imbalance feedback procedure. Lastly, FAST presents only one stimulus at each trial, therefore, it is more difficult for participants to predict what is being assessed during the FAST task. For such reasons, FAST might be useful for investigating the defusion process.

(4) A Potential Analytical Method for the FAST: The Differential Arbitrarily Applicable Relational Responding Effect (DAARRE) Model

The defusion process, which reduces the believability of self-statements, can be assessed by

FAST. Given that the negative and positive self-statements may be arbitrarily related to stimuli that have similar functions, the FAST consistent block may ask participants to establish negative self-statement — aversive stimuli and positive self-statement — appetitive stimuli in the consistent block. In this context, participants were expected to acquire the response class more easily in the consistent block than in the inconsistent block. Nevertheless, the likelihood of AARR emergence varied across people. Specific AARR is more likely to emerge under certain conditions.

A study recently proposed a model called the differential arbitrarily applicable relational responding effect (DAARRE) model, that demonstrated its usefulness in analyzing and predicting the emergence probability of AARR under the laboratory setting at individual and group levels in terms of the related stimulus functions and properties of stimulus relation (e.g., see Finn, Barnes-Holmes, and McEnteggart 2018). According to the model, a produced response under the laboratory setting is the result of the dynamic interaction among three factors: (1) a context that controls and informs how stimuli are related (Crel) such as coherence (i.e., the degree of functional coherence between stimuli in a relation), complexity (the density or quality of AARR such as number of stimuli involves and type of AARR), derivation (i.e., the extent to which a certain AARR has been practiced or emerged) and flexibility(the extent to which a certain AARR could be altered in respect to the context); (2) a context that determines the transformation that occurs in a stimulus function and how it takes place (Cfunc), and (3) relational coherence indicators (RCIs) indicating whether a presented stimulus relation is coherent or incoherent to one's learning history (Pidgeon et al. 2021; Finn, Barnes-Holmes, and McEnteggart 2018). The properties of these factors are indicated by a − or + sign, and the response that is most consistent in terms of the sign is thought to be the one that would most likely emerge.

To illustrate how the DAARRE model can be used for the FAST hypothesis, it is important to determine Cfunc, Crel, and RCIs involved in the FAST. For example, FAST in the current study to access the defusion process of the defusion procedure such as LYT asked participants to establish the response class of (1) negative self-statements – "might be false" stimulus that evokes negative feelings and (2) positive self-statements – "might be true" stimulus that evokes positive feelings. "Might be false/ true" were selected as FAST stimuli because they indicate the degree of coherence in a context, and being coherent is reinforcing for humans which is learned via social interaction (Villate et al. 2016). Therefore, "might be false" was assumed to work as an aversive stimulus that evokes negative feeling, and "might be true" were assumed to be an appetitive stimulus that evokes positive feelings.

In this FAST, the following properties of Cfunc, Crel, and RCIs should be considered (see Figure 1). First, there are four Cfunc for the presented stimuli during the task. The Cfunc for "I am a bad person" (e.g., "I am bad") evokes an aversive feeling compared with the Cfunc for "I am a good person," which suggests an appetitive feeling. These are denoted as Cfunc − (a relatively aversive evoking function) and Cfunc + (a relatively appetitive evoking function). In addition, the Cfunc for "might be false" indicates an aversive feeling since "false" indicates an incoherence of facts that is punished via one's social learning compared with "might be true," which has a Cfunc evoking an appetitive feeling. These are also denoted as Cfunc − and Cfunc +, respectively. Second, the FAST involves four Crels with two aspects each. For instance, "I am a bad person" and "might be false" are functionally coherent (i.e., Cfuncs match each other) and low in complexity because the AARR in this context only includes coordinated relations, the simplest stimulus relation. These are indicated as Crel coherence + and Crel complexity −, respectively. The complexity was implemented for hypothesis formation; LYT is the manipulation of context where original self-statements were added into hierarchical relation of "I am having a thought that…," which could be considered as manipulating the complexity of self-statements related AARR.

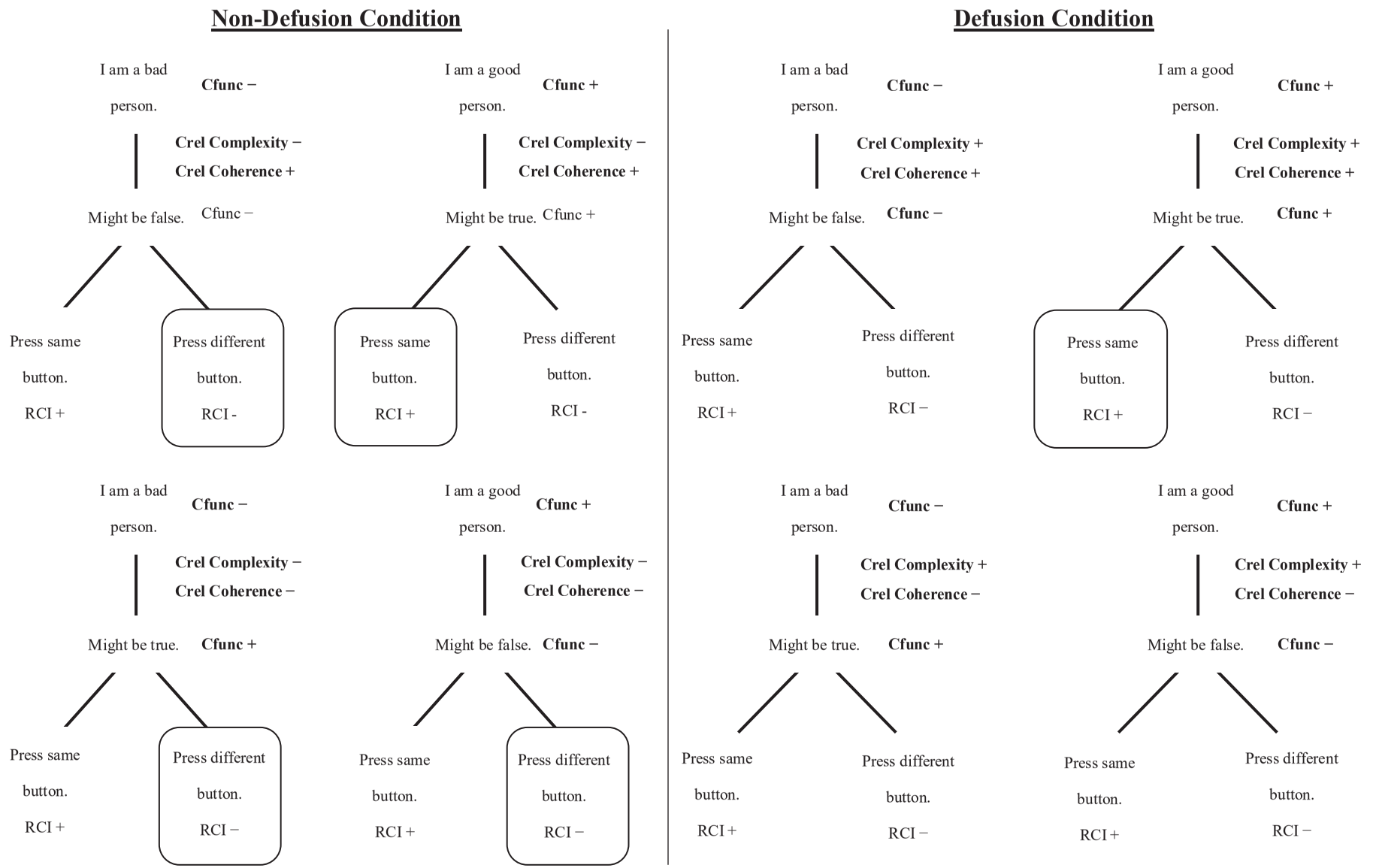

Figure 1. Response Pattern of a Participant in FAST

Source: Author

Note: The enclosed button-pressing response with the rectangle is the one that is more likely to be provided than the other response. The four response patterns on the left are the non-defusion condition, and the other four response patterns on the right side are the defusion condition.

"I am a good person" and "might be false" are functionally incoherent and less complex. These are indicated as Crel coherence − and Crel complexity −. Meanwhile, "I am a bad person" and "might be true" are not functionally coherent (Crel coherence −) and are low in complexity (Crel complexity −). Also, the Crels for "I am a good person" and "might be true" are functionally coherent (Crel coherence +) and have low complexity (Crel complexity −). RCIs are unique in the case of the FAST, which involves pressing the functionally coherent keys M or Z. Put simply, when participants provide the correct response to a set of stimuli, this indicates the functional coherence of that stimuli set (RCI +) whereas an incorrect response refers to its functional incoherence (RCI -).

The DAARRE model suggested that the AARR with the greatest number of + or – is most likely to occur. Thus, individuals having a thought "I am a bad person" and "I am a good person" respond in such a way that they obtain the + or − sign the most in FAST. It is equally difficult for them to relate "I am a bad person" with "might be false" the most (four – vs. two +), "I am a bad person" with "might be true" (four – vs. two +), and "I am a good person" with "might be false" (four – vs. two +). In addition, relating "I am a good person" and "might be true" (two – vs. four +) is the easiest for the individuals. Accordingly, they are expected to outperform in the consistent block (i.e., relating "I am a bad person"–"might be false" and "I am a good person"–"might be true" are required by pressing same button) over the inconsistent block (relating "I am a bad person"–"might be true" and "I am a good person"–"might be true" are required by pressing same button), as relating "I am a good person" with "might be true" is the easiest whereas remaining relations are equally difficult to be learned. Therefore, FAST SoR for the participant would be a positive number and significantly different from 0.

However, an individual using defusion, LYT, for instance, might produce a slightly different response for FAST. Although the Cfunc and RCIs would not change, Crel would be altered in terms of complexity. LYT makes individuals recontextualize their self-statements in the form of "I am having a thought that [a self-statement]." Compared to original self-statements, adding prefix of LYT led increasing number of stimuli involved in the self-statement relation, thus, they were required to relate "I am a bad/good person" with "might be true/false" while relating "I am having a thought," which in turn increases relational complexity as the density of AARR is greater and the more complex AARR type, the hierarchical relation (i.e., "I am having a thought…") is involved (see Figure 1). As a result, relating "I am a bad person" with "might be false" (three – vs. three +), "I am a bad person" –with "might be true" (three – vs. three +), "I am a good person" with "might be false" (three – vs. three +) are equally difficult; relating "I am a good person" with "might be true" (one – vs. five +) is the easiest to be related as stimuli because their relation showed the maximum coherence of +. That is, the individuals also find it easier to relate "I am a good person" with "might be true" even in comparison with a situation where they did not use LYT. Thus, the individuals using LYT may be better at learning the stimulus relation in the consistent block than in the inconsistent block, resulting in greater SoR than 0 and SoR for those who did not use LYT. Nevertheless, the hypotheses were not tested.

(4) Objective

This study aims to investigate the defusion process of "LYT" by using FAST, in which their reported two positive and negative self-statements in the questionnaire as well as the phrases "might be true," "might not be false," "might be false," and "might not be true" were presented as experimental stimuli. The psychotherapist frequently rephrases clients' rigid and inflexible thought by saying "you are having a thought that…" in order to encourage the sense that the thought is not absolute truth (e.g., Hyddy et al. 2016). Therefore, using LYT in the study, which allows participants

to rephrase their self-statements in the form of "I am having a thought that…" may contribute to gaining insights about how clinical conversations can affect clients although it is not the main point of this study. Participants using LYT may learn the stimulus class in the consistent block easier than in the inconsistent block, and participants not using LYT may perform equally in both blocks. Those using LYT may display a greater SoR than those not using LYT.

At the same time, the defusion outcome of LYT was assessed to ensure that LYT works sufficiently. First, the believability of positive and negative self-statements is compared between individuals using LYT and not using LYT. Healy et al. (2008) suggested that LYT had only limited effect in the context of positive self-statement compared to the context of negative self-statement. Therefore, it was hypothesized that LYT decreases the believability of negative self-statements but not positive self-statements. Subsequently, the study examines whether there are differences in the correlation between the believability of positive and negative self-statements among individuals using and not using LYT. Accepting or learning self-statements that contradict one's learning history is challenging. For instance, a participant who strongly believes that "I am a bad person" finds it difficult to learn a positive self-statement (e.g., "I am a good person") since the participant has already learned that he or she was defined as "a bad person," contradicting the relation "I" am equivalent to "a good person." In other words, the believability of positive self-statements is expected to be negatively correlated with the believability of negative self-statements. However, there may not be a correlation between the believability of positive and negative self-statements in the context of LYT use due to the recontextualization of one's self-statements in a hierarchical relation (e.g., "I am having a thought that XXX"). Informally speaking, individuals using LYT obtain space for multiple self-statements by viewing their thoughts as a part of themselves. Thus, it is considered that the believability of positive and negative self-statements for the individuals using LYT is not correlated. In summary, it was hypothesized that participants engaged in LYT are expected to report lower believability of negative self-statements and no correlation between the believability of negative and positive self-statements. Conversely, those who did not engage in LYT may report higher believability of negative self-statements and a negative correlation between the believability of negative and positive self-statements.

2. Method

(1) Participants

A total of 23 individuals (male = 4, female = 19; mean age = 26.52, age *SD* = 6.04) who were Japanese native speakers or non-Japanese native speakers with an N1 Japanese-language proficiency participated in this study and were randomly divided into two groups. A total of 10 participants (male = 3, female = 7; mean age = 26.5, age *SD* = 5.69) were assigned to the defusion condition, and the remaining were assigned to the non-defusion condition (male = 1, female = 11; mean age = 26.5, age *SD* = 6.29).

(2) Instruments

The participants answered an information sheet, two self-report questionnaires (a self-statement sheet and a believability sheet), a defusion worksheet, and the FAST. The participant information sheet included the participant's name, gender, age, and contact information. The self-statement sheet required participants to write down two negative self-statements (N1 and N2) and two positive self-statements (P1 and P2) starting with "I," such as "I am worthless." These

statements were used to create the believability sheet as well as the stimuli presented in the FAST. The believability sheet asked participants to rate the believability of statements N1, N2, P1, and P2 from 1 (not believe) to 100 (true). The believability sheet for the participants in defusion condition included the instruction "CAUTION: Please use what you have learned from the defusion worksheet while filling out this questionnaire" in red text to ensure they used the defusion process from the defusion worksheet. The defusion worksheet asked participants to write down N1, N2, P1, and P2 again and then rewrite each self-statement in the form of "I am having a thought that <statement>" (e.g., "I am having a thought that I am worthless."). At the end of the worksheet, the participants were requested to practice defusion worksheet on their self-statements or thoughts about themselves since defusion is a skill to be practiced and used fluently whenever the need for it arises. However, they were not explicitly instructed in the way of practicing the defusion worksheet (i.e., they were free to use defusion worksheet, but it was not required). Meanwhile, FAST is a computer-based assessment tool to assess stimulus relatedness. This study used a laptop to administer FAST 2018 v. 1 (Roche and Stewart 2018). The participants sat in front of the laptop, were presented with a stimulus on the screen during the FAST and were required to respond to the stimulus by pressing Z or M on the keyboard. All these responses were recorded using the FAST software.

(3) General Procedure

The participants filled out the participant information sheet and the self-statement sheet. For the self-statement sheet, the participants wrote down two negative and two positive self-statements. Those assigned to the defusion condition were asked to complete the defusion worksheet. All participants chose the schedule for their second experimental day, which must be at least one week after the first experimental day. The participants in the defusion condition were asked to practice the defusion worksheet until their next experimental day.

On their second experimental day, the participants took the FAST, which was divided into two parts: one for practice and the other for correcting experimental data. After completing the FAST, the participants in the defusion condition filled out a believability sheet and completed a Google Form questionnaire inquiring about the number of times they practiced the defusion worksheet. At the end of the current study, the non-defusion participants were informed that they were in the non-defusion condition; they listened to an instruction video for the defusion worksheet given to the defusion condition, and the defusion worksheet was e-mailed to them for the purpose of ethical consideration.

(4) FAST

During the FAST, the participants were presented with a random stimulus with a black color at the center of the computer screen and responded to it by pressing the Z or M key. Each stimulus was presented on the screen for 3,000 ms. After the participants pressed Z or M, the message "Correct (i.e., 正解)" or "Incorrect (i.e., 不正解)" in red text appeared on the screen, and the next trial began. Based on the message, the participants needed to learn how to correctly respond to stimuli. In addition, if they did not respond within 3,000 ms, they received feedback of "Incorrect." At the beginning of each block, the following instruction was given: "In this section, letters will appear on the screen. You will learn which keys should be pressed in the presence of the letters on the screen. Caution: in this phase, only Z key, and M key will be used. The session will continue until you get 10 consecutive correct responses. Press the enter key to start this session. When you press the key, the section will start." To better understand the FAST process, explanations were provided for the

consistent and inconsistent blocks, followed by baselines 1, and 2.

a) Consistent and Inconsistent Blocks

The presentation sequence of the inconsistent and consistent blocks was randomized by software. In the consistent block for the practice FAST, if nonsense syllable stimuli A3, B3, C3, or D3 were presented, the participants were required to press the M key; if nonsense syllable stimuli A4, B4, C4, or D4 appeared on the screen, they had to press Z. Meanwhile, in the experimental FAST, a participant's N1 and N2, "might be false," or "might not be true" was presented, and the participant had to press M in this case; if the participant's P1 and P2, "might be true," or "might not be false" was displayed on the screen, they had to press Z.

In the inconsistent block for the practice FAST, the participants had to press M when nonsense syllable stimuli A3, B3, C4, or D4 appeared on the screen and Z if nonsense syllable stimuli A4, B4, C3, or D3 were displayed. In the experimental FAST, the participants needed to press M for N1, N2, "might be true," or "might not be false" and Z for P1, P2, "might be false," or "might be not true."

b) Baselines 1 and 2

For baseline 1 of the practice FAST, the participants had to press Z when nonsense syllable stimuli A1, B1, C1, or D1 were displayed and M key when nonsense syllable stimuli A2, B2, C2, or D2 appeared. For baseline 2 of the practice FAST, the participants needed to press Z when nonsense syllable stimuli A5, B5, C5, or D5 were shown, and M when nonsense syllable stimuli A6, B6, C6, or D6 were displayed. For baseline 1 of the experimental FAST, the participants were required to press M for alf, wnf, nvq, or mip and Z for lnj, twr, gbl, or jhs. For baseline 2 of the experimental FAST, the participants needed to press M for lnj, twr, gbl, or jhs and Z for dlp, eqs, pnf, or hkf.

(5) Analysis Data

SoR, an index of the FAST, had a criterion in which participants were required to produce 10 consecutive correct responses within 100 trials in each block, and those who could not reach to the criterion would be excluded from all data analysis. SoR is obtained by subtracting the number of trials to reach 10 consecutive correct responses in the consistent block from that in the inconsistent block and dividing this number by the average number of trials to each 10 consecutive correct responses for baselines 1 and 2. A positive score indicates the ease of learning stimulus relations in the consistent block, while a negative score pertains to the ease of learning stimulus relations in the inconsistent block.

(6) Ethics Consideration

This study obtained approval from the ethics committee of the university where the author worked. It ensured the protection of human rights (e.g., the right to participate and leave the study, data security) and well-being of participants (e.g., the participant's mental and physical safety).

3. Results

Of the 10 participants assigned to the defusion condition, five practiced LYT every day, one practiced once in two days, and four practiced once in three days. Table 1 shows the participants' strength of relation (SoR) and the believability rates for their positive and negative self-statements.

Table 1. Number of Trials Required to Meet the FAST Criterion, SoR, and the Mean Believability of Negative and Positive Self-Statements

Participants	001	002	003	004	005	006	007	008	009	010	011	012	013	014	015	016	017	018	019	020	021	022	023
Condition	1	1	2	1	2	2	1	2	2	1	1	1	2	2	2	2	2	1	1	2	2	2	1
Baseline 1	over	41	34	over	over	56	43	-	31	40	37	31	54	58	85	22	16	97	28	over	64	25	35
Consistent	66	17	12	over	22	26	79	-	11	66	20	23	39	23	54	42	13	16	17	over	18	53	22
Inconsistent	26	38	31	62	39	24	26	-	14	over	12	20	over	44	13	12	10	31	61	over	23	37	39
Baseline 2	83	59	32	59	40	23	55	-	35	46	49	17	19	78	35	37	13	over	18	over	34	40	26
SoR	-	0.42	0.58	-	-	-0.05	-1.08	-	0.09	-	-0.19	-0.13	-	0.31	-0.68	-1.02	-0.21	-	1.91	-	0.10	-0.49	0.56
Negative Self-Statement Believability	60	50	100	80	85	35	27.5	72.5	67.5	64	90	37.5	72.5	75	87.5	90	77.5	40	35	80	70	100	90
Positive Self-Statement Believability	35	70	100	65	95	50	85	92.5	67.5	55	70	72.5	90	90	100	95	70	70	67.5	90	80	90	90

Note: Condition 1 is defusion and 2 is non-defusion. Baseline 1, 2, Consistent, Inconsistent refer to the number of required trials to reach FAST criterion in the *block*. "Over" means that the participants could not reach to the FAST criterion in the block, and "-" means no data. If a participant's data had either "over" or "-", they were excluded from the data analysis.

The Mann-Whitney U test was performed to assess whether SoR for the defusion condition was greater than non-defusion. The SoR in defusion condition (Md = 0.14) was not significantly greater than SoR in non-defusion condition (Md = 0.05, Z = - 0.94, p = .388), indicating that the hypothesized process of defusion of LYT was unsupported.

Next, the difference between the means believability of positive and negative self-statements between conditions and their correlational differences was examined to see if LYT works successfully or not. First, the mean believability scores for the positive and negative self-statements were calculated and analyzed using the Mann-Whitney U test. Results showed the mean believability of a positive self-statement in defusion condition (Md = 71.25) was not significantly different from the mean believability of a positive self-statement in a non-defusion condition (Md = 90, Z = - 0.94, p = .38). There was also no significant difference in the mean believability of a negative self-statement between defusion (Md = 43.75) and non-defusion conditions (Md = 77.50, Z = 1.30, p = .22). Second, two-tailed Spearman's rho was conducted to see the correlation between negative and positive self-statements in defusion and non-defusion conditions. Although there was no correlation between the mean believability of negative and positive self-statements in the defusion condition (r (6) = .13, N = 6, p = .80), there was a significant correlation between the mean believability of negative and positive self-statements in the non-defusion condition (r (9) = .81, N = 9, p = .008).

4. Discussion

This study investigated the defusion process of LYT in changes surrounding the believability of positive and negative self-statements by using FAST and DAARRE models. No changes or processes were found in the behavioral index assessed by the FAST as there were no significant differences in SoR between the defusion and non-defusion conditions. Furthermore, a positive correlation, instead of a negative correlation, between the believability of positive and negative self-statements was only observed in the non-defusion condition, and the result showed that LYT did not work as there was no significant difference in the mean believability of positive and negative self-

statements between the two conditions.

The potential reason for the non-significant result of the FAST was the failure to capture the unstable effect of LYT due to experimental procedure issues. The defusion procedure, including LYT, disrupts the process of verbal behaviors, but it must be constantly practiced to allow stable changes in the process of verbal behaviors. However, the current study did not control the degree of proficiency in utilizing LYT. Therefore, the participants' proficiency level at utilizing LYT might not be enough to yield stable effects or changes, thereby, no significant result was observed by FAST. In addition, the greater stimulus relatedness due to a greater number of learning yields the greater FAST effect size (Cummins et al. 2018). In the current study, the number of LYT practices was not controlled; thus, they varied across participants. For those reasons, FAST failed to capture the process of defusion of LYT and its effect. The other procedural issue was that participants in the defusion condition were asked to use LYT not only before rating the believability of self-statements, but also before FAST. Even though the LYT effect was insufficient, this issue might also affect the FAST effect.

The second potential reason for the FAST result is that the variability of the Cfunc of a participant's self-statement derived from their preexisting history might affect the SoR variability. For example, participants who strongly believe negative self-statements may experience more intense aversive functions for such statements than other participants. It is well known that pre-existing history related to the emergence of AARR affects FAST scores The Cfunc of stimuli in an experimental procedure such as participants' self-statements could dynamically interact with properties of Crel, Cfunc of other stimuli, and RCIs, which in turn may produce a wide variability in FAST scores. Nonetheless, the study did not assess the believability or original SoR in participants' self-statements before intervention; hence, it could not be investigated. Also, the stimuli "might be true/false" might fail to evoke the targeted Cfunc expected, resulting in the rejection of the hypothesis. The current study selected the word "might be true" as the appetitive stimulus (Cfunc +) compared to the "might be false" as the aversive stimulus (Cfunc −), however, relating a positive self-statement to "might be false" for instance is easier for the participants who strongly believe the negative self-statements because it confirms the positive self-statement was not true. "Might be false" works as Cfunc +. Indeed, the Japanese language is highly context-dependent language (Hirose and Hasegawa 2010), meaning that the Cfunc of verbal stimuli highly depends on the context, therefore, it is possible that "might be true" and "might be false" failed to evoke the Cfuncs that were anticipated in the current study.

The last potential reason for the FAST result is the unexpected dynamic interaction among DAARRE factors. In the current study, LYT was assumed to affect coherence and complexity by recontextualizing self-statements with the prefix of "I am having the thought that…." However, the changes in Cfuncs, derivation, and flexibility were not anticipated. In fact, the DAARRE model argues that all the factors in the model interact dynamically. It is possible to say that LYT may affect Cfunc directly, or LYT may affect complexity that alters the derivation and flexibility for instance. However, the DAARRE model is newly developed, and there is not sufficient data yet to analyze which factor was affected by LYT and how changes in one factor can affect the others. Thus, more accumulation of data using DAARRE model would be needed.

The current study has four limitations. First, the experiment was not pre-post design, leading to the failure of assessing the Cfuncs of participant's original self-statements. Second, the "might be true/false" stimuli were used and potentially evoked unexpected Cfuncs as discussed above. Third, the degree of proficiency in utilizing LYT was not controlled across all participants. Last, FAST was administered before the believability questionnaire which they filled out while only defusion participants were asked to use what they have learned by the defusion worksheet. These

four limitations could lead to an unpredicted result for LYT and FAST. Therefore, future studies need to employ a pre-post experimental design with multiple conditions which differ in the degree of proficiency at utilizing LYT (e.g., single-practice LYT condition vs. 10-practice LYT condition) and involve FAST presenting self-statements and stimuli that evoke stronger aversive or appetitive functions like "disgusting" and "loving" instead of "False" or "True." Additionally, an instruction like "please try to work on the task while you use what you have learned during experiments for the self-statements" before and after the FAST could reduce the variability of LYT effect.

This study examined the LYT process by focusing on Crel, Cfunc, and RCIs by FAST and DAARRE Models. Although the defusion-the-process remains unclear, it highlights that a more detailed analysis of Cfunc can contribute to this type of experiment. The self-statements are perhaps the most complex and well-practiced AARR that is not easily changed but heavily influences one's behaviors (McHugh, Stewart, and Priscilla 2019). Recently, more detailed analytical models of AARR are now available (see Harte and Barnes-Holmes 2021). This might be contributable to combining the DAARRE model and the FAST experiment, however, further experiments are needed to consider its utility in the defusion process and phenomenon related to self-statements. Understanding self-statements is critical since it sometimes aversively impacts an individual's flexibility to behave in accordance with the present moment. Psychotherapy now implements an analytical method focusing on stimulus relation and function, called verbal functional analysis (Barnes-Holmes et al. 2018). Verbal functional analysis is extremely effective for case formulation as it can help analyze the history of a client's problem or difficulty and therefore aid in the development of a plan to approach the client in that situation. It enhances one's understanding of the client's situation in terms of AARR and allows therapists and clients to recontextualize in a way that neither avoids nor denies existing AARR. To conclude, the current study was unable to reveal the defusion process. However, it has become one of the few studies that have attempted to investigate the defusion process by using FAST and DAARRE models. Although further studies are definitely required to investigate the defusion process, this study helps strengthen the background knowledge for investigating the defusion process and opens opportunities for future exploration.

References

Assaz, Daniel A., Bryan Roche, Jonathan W. Kanter, and Claudia K. B. Oshiro. 2018. Cognitive Defusion in Acceptance and Commitment Therapy: What Are the Basic Processes of Change? *The Psychological Record*, 68(4), 405–418. <https://doi.org/10.1007/s40732-017-0254-z>

Atkins, Paul W. B., and Robert Styles. 2015. Mindfulness, Identity and Work: Mindfulness Training Creates a More Flexible Sense of Self. In Jochen Reb, and Paul W. B. Atkins (eds.), *Mindfulness in Organization: Foundations, Research, and Applications*. Cambridge: Cambridge University Press, 133–162. <https://doi.org/10.1017/CBO9781107587793.008>

Barnes-Holmes, Yvonne, John Boorman, Joseph E. Oliver, Miles Thompson, Ciara Mcenteggart, and Carlton Coulter. 2018. Using Conceptual Developments in RFT to Direct Case Formulation and Clinical Intervention: Two Case Summaries. *Journal of Contextual Behavioral Science*, 7, 89–96. <https://doi.org/10.1016/j.jcbs.2017.11.005>

Beadman, Matthew, Ravi K. Das, Tom P. Freeman, Peter Scragg, Robert West, and Sunjeev K. Kamboj. 2015. A Comparison of Emotion Regulations Strategies in Response to Craving Cognitions: Effects on Smoking Behaviour, Craving and Affect in Dependent Smokers. *Behaviour Research and Therapy*, 69, 29–39. <https://doi.org/10.1016/j.brat.2015.03.013>

Blackledge, John T. 2007. Disrupting Verbal Processes: Cognitive Defusion in Acceptance and Commitment Therapy and Other Mindfulness-based Psychotherapies. *The Psychological Record*, 57(4), 555–576. <https://doi.org/10.1007/BF03395595>

Cummins, Jamie, Bryan Roche, Ian Tyndall and Aoife Cartwright. 2018. The Relationship Between Differential Stimulus Relatedness and Implicit Measure Effect Sizes. *Journal of the Experimental Analysis of Behavior*, 110(1), 24–38. <https://doi.org/10.1002/jeab.437>

Finn, Martin, Dermot Barnes-Holmes, and Ciara McEnteggart. 2018. Exploring the Single-Trial-Type-Dominance-Effect in the IRAP: Developing a Differential Arbitrarily Applicable Relational Responding Effects (DAARRE) Model. *The Psychological Record*, 68(1), 11–25. <https://doi.org/10.1007/s40732-017-0262-z>

Finn, Martin, Dermot Barnes-Holmes, Ciara McEnteggart, and Deirdre Kavanagh. 2019. Predicting and Influencing the Single-Trial-Type-Dominance-Effect: the First Study. *The Psychological Record*, 69, 425–435. <https://doi.org/10.1007/s40732-019-00347-4>

Harte, Colin, and Dermot Barnes-Holmes. 2021. Whatever I "ROE-M" there I am: An RFT (technical) Account of the Verbal Self and Altered Stages of Consciousness. Association for Behavior Analysis International. <https://science.abainternational.org/2021/03/29/wherever-i-roe-m-there-i-am-an-rft-technical-account-of-the-verbal-self-and-altered-states-of-consciousness/>

Healy, Hilary-Anne, Yvonne Barnes-Holmes, Dermot Barnes-Holmes, Claire Keogh, Carmen Luciano, and Kelly Wilson. 2008. An Experimental Test of a Cognitive Defusion Exercise: Coping with Negative and Positive Self-statements. *The Psychological Record*, 58(4), 623–640. <https://doi.org/10.1007/BF03395641>

Hirose, Yukio, and Yoko Hasegawa. 2010. *Nihongo Kara Mita Nihonjin: Shutaisei no Gengogaku* (*Japanese People seen Through a Japanese Language: Linguistic Study of Subjectivity*). Tokyo: Kaitakusha. (In Japanese).

Hooper, Nic, and Louise McHugh. 2013. Cognitive Defusion Versus Thought Distruction in the Mitigation of Learned Helplessness. 2013. *The Psychological Record*, 63(1), 209–218. <https://doi.org/10.11133/j.tpr.2013.63.1.016>

Hyddy, Vyv, Anna Roberts, Manulea Jarrett, and Lucia Valmaggia. 2016. Psychological Therapy for At Risk Mental State for Psychosis in a Prison Setting: A Case Study. *Journal of Clinical Psychology*, 72(2) 142-151. <https://doi.org/10.1002/jclp.22250>

Masuda, Akihiko, Steven C. Hayes, Casey F. Sackett, and Michael P. Twohig. 2004. Cognitive Defusion and Self-Relevant Negative Thoughts: Examining the Impact of a Ninety Year Old Technique. *Behaviour Research and Therapy*, 42(4), 477–485. <https://doi.org/10.1016/j.brat.2003.10.008>

McHugh, Louise, Ian Stewart, and Priscilla Almada. 2019. *A Contextual Behavioral Guide to the Self: Theory and Practice*. California: New Harbinger Publications.

O'Reilly, Anthony, Bryan Roche, Maria Ruiz, Ian Tyndall, and Amanda Gavin. 2012. The Function Acquisition Speed Test (FAST): A Behavior Analytic Implicit Test for Assessing Stimulus Relations. *The Psychological Record*, 62(3), 507–528. <https://doi.org/10.1007/BF03395817>

Pidgeon, Audrey, Ciara McEnteggart, Colin Harte, Dermot Barnes-Holmes, and Yvonne Barnes-Holmes. 2021. Four Self-related IRAPs: Analyzing and Interpreting Effects in Light of the DAARRE Model. *The Psychological Record*, 71(3), 397–409. <https:/doi.org/10.1007/s40732-020-00428-9>

Roche, Bryan, and Ian Stewart. 2018. Function Acquisition Speed Test (FAST) 2018 v.1. Computer Software.

Society of Clinical Psychology. 2022. Psychological Treatments. Society of Clinical Psychology. <https://div12.org/treatments>

Villatte, Matthieu, Jennnifer L Villatte, and Steven Hayes. 2016. *Mastering the Clinical Conversation: Language as Intervention*. New York: Guilford Press.