

From Felt Sense to Words: Construction and Analysis of a Focusing Dialogue Dataset for Verbalization

Yuiko Tsunomori[†], Yuya Chiba[†], and Yosuke Koshikawa[◇]

[†]NTT, Inc.

[◇]Kansai Medical University

{yuiko.tsunomori,yuya.chiba}@ntt.com, koshikawa.you@kmu.ac.jp

Abstract

Verbalization, the process of expressing one's internal states in words, has been shown to deepen self-understanding and improve well-being. However, this can be difficult to achieve for some individuals. One approach to supporting verbalization is *Focusing-Oriented Psychotherapy*. Focusing facilitates verbalization by guiding a speaker's attention to their *felt sense*, an internal state that has not yet been verbalized, and helping them to find the words that appropriately express this felt sense through dialogue. Toward developing Focusing dialogue systems that support user verbalization, we constructed a dataset containing Focusing dialogues between professionally trained listeners and speakers. This dataset consists of 50 dialogues (762 minutes, 17,986 utterances) and contains ratings of verbalization progress, subjective evaluations, and dialogue act (DA) annotations. To clarify the characteristics of Focusing dialogues, we analyzed associations among verbalization progress, subjective evaluations, and DA usage. Based on these analyses, we discussed design implications for Focusing dialogue systems.

1 Introduction

Verbalization is defined as expressing one's internal states—thoughts, feelings, and fantasies—in words¹. Verbalization not only makes internal states shareable with others, but also contributes to deeper self-understanding and improvements in health-related outcomes (Pennebaker and Chung, 2000; Ozomaro et al., 2025). However, some individuals experience difficulty in producing verbalization (Barrett et al., 2001). This inability highlights the importance of developing means to support verbalization.

One approach to supporting verbalization is *Focusing-Oriented Psychotherapy* (hereafter, Fo-

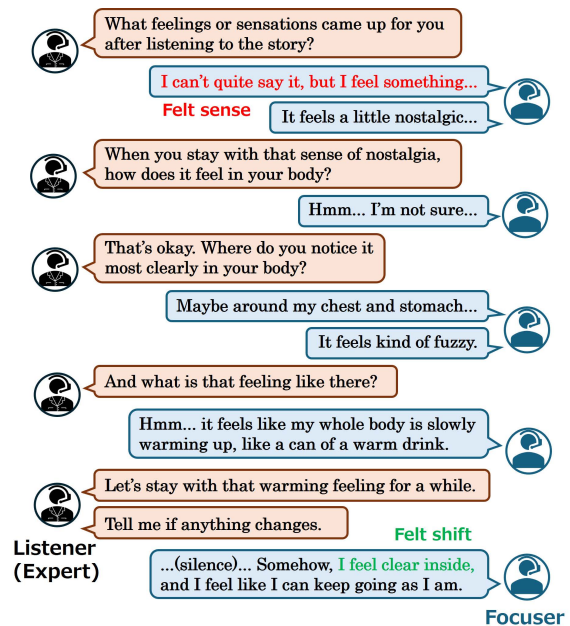


Figure 1: Example of a Focusing dialogue between a listener and a focuser.

cusing) (Gendlin, 1981). Focusing, developed through counseling research, is a process of attending to an internal state that has not yet been verbalized and gradually putting it into words, thereby facilitating psychological change (Hendricks, 2002). Here, a pre-verbal internal state is called a *felt sense*, and the change that occurs when the words match the felt sense is called a *felt shift*. In Focusing, the person who engages in this process is called the *focuser*. Although focusers can practice Focusing alone, they usually find it more effective with the support of a professionally trained *listener*. Figure 1 shows an example of a Focusing dialogue, where the focuser gradually clarifies their felt sense with the listener's support and thus achieves a felt shift. Unfortunately, the availability of trained listeners is limited, which makes it difficult to carry out effective Focusing in a practical way. This motivates the

¹<https://dictionary.apa.org/verbalization>

use of automated dialogue systems for practicing Focusing, where the system serves as a listener to make verbalization support more accessible in everyday settings.

To develop Focusing dialogue systems, it is necessary to first understand the characteristics of Focusing dialogues. Although Focusing is often described as a structured process (Gendlin, 1981), listeners in actual Focusing dialogues adapt their responses depending on the focuser's state. Therefore, process descriptions alone cannot sufficiently capture how listeners support the focuser's verbalization in practice. Furthermore, the effectiveness of Focusing can be assessed in terms of both *verbalization progress* and the speaker's *subjective evaluations* (Lietaer and Neirinck, 1986), such as perceived changes and satisfaction, but it remains unclear how the two are related. Understanding these points is important in the design and evaluation of Focusing dialogue systems. Such understanding requires datasets of actual Focusing dialogues, but to the best of our knowledge, no publicly available dataset of sufficient scale exists.

In this study, we constructed a multimodal Focusing dialogue dataset. Since Focusing is centered on the focuser's bodily sensations, non-verbal information is also important. Therefore, we collected audio and video recordings of Focusing dialogues that capture speakers' bodily, vocal, and facial expressions. The collected dataset consists of 50 Focusing dialogues (762 minutes and 17,986 utterances) between professionally trained listeners and focusers. Each dialogue contains ratings of verbalization progress, subjective evaluations by both listeners and focusers, and dialogue act (DA) (Stolcke et al., 2000) annotations. We then analyzed the collected dataset from two perspectives. Specifically, we examined how verbalization progress is associated with speakers' subjective evaluations and analyzed utterance patterns characteristic of Focusing. Finally, based on these analyses, we discussed design implications for Focusing dialogue systems.

2 Focusing-Oriented Psychotherapy

Focusing has been proposed as an approach to psychotherapy based on findings in counseling research (Gendlin, 1961, 1981, 1996). Although Focusing is based partly on Person-Centered Therapy (PCT), which is known as one of the three major approaches to psychotherapy (Rogers, 1951,

1961; Raskin and Rogers, 2005), it is characterized by a more explicit emphasis on the *bodily felt sense* (Ikemi, 2005).

2.1 Focusing in Practice

In general, the Focusing process is formulated as a sequence of the following six steps (Gendlin, 1981). (1) *Clearing a Space*: attending to internal states in the body. (2) *Felt Sense*: selecting a specific aspect of one's internal states to focus on. (3) *Handle*: finding the words that clearly express this felt sense. (4) *Resonating*: checking how the words will match the felt sense. (5) *Asking*: exploring possibilities of more suitable words. (6) *Receiving*: welcoming whatever comes with the felt shift. Listeners do not necessarily follow these steps exactly but instead adapt flexibly to the focuser's state and the flow of the dialogue.

In Focusing, focusers may encounter various difficulties, such as failing to recognize the felt sense or being unsure where to direct their attention. To make Focusing more approachable, especially for beginners, *Focusing works* with specific themes have been proposed. In *Focusing-Oriented Dreamwork* (Ellis, 2014), the focuser recalls a dream they had in sleep and practices Focusing based on this recollection. In *Crossing with Animals* (Ikemi, 2024), the focuser expresses their felt sense through animal metaphors. These kinds of Focusing works provide the focuser with concrete materials, such as images and metaphors, that serve as an entry point for attending to the felt sense.

2.2 Measures of Focusing

As a measure for assessing the effectiveness of Focusing, the Experiencing Scale (EXP scale) (Klein et al., 1969) is often employed. It is used by the listener to evaluate verbalization progress for each utterance, namely how well the focuser puts the felt sense into words. However, applying the EXP scale is costly because it requires detailed utterance-level annotation. To address this issue, an EXP checklist (Ohashi and Ikemi, 2017) has been proposed, in which the listener evaluates verbalization progress throughout the dialogue. The EXP checklist consists of 17 items for estimating the dialogue-level EXP scale, each classified into one of five evaluative levels ranging from Very Low to Very High (mapped to scores from 1 to 5). The listener determines whether each item applies to the dialogue, and the highest level among those

corresponding to the applicable items is adopted as the final *EXP score*. This checklist is shown in Table 9 in the Appendix.

The effectiveness of Focusing can be influenced by the focuser’s pre-existing attitude toward attending to the felt sense (Kiesler, 1971). To assess this attitude, the Focusing Manner Scale (FMS) (Aoki and Ikemi, 2014) was proposed, and its revised version, FMS-18 (Morikawa et al., 2014; Fukumori, 2022), has been widely used. This is an 18-item scale that measures how individuals deal with their felt sense in daily life, and these items are categorized into three factors: *attention*, paying attention to their felt sense; *distance*, keeping some distance from it; and *acceptance*, accepting it as it is. Table 10 in Appendix lists all items in the FMS-18.

3 Related Work

3.1 Analysis of Focusing

Previous studies have primarily focused on the effectiveness of Focusing in clinical counseling settings. In these studies, the EXP scale (Klein et al., 1969) has been widely used, and it is associated with positive changes such as self-compassion and self-understanding (Kiesler, 1971; Hendricks, 2002; Pascual-Leone and Yeryomenko, 2017; Gomes et al., 2024; Wong, 2025). Further studies have investigated focusers’ pre-existing Focusing attitudes using FMS (Aoki and Ikemi, 2014), and FMS scores have been associated with stress responses, psychological well-being, and self-confidence (Yoshiwara and Tsuchiya, 2019; Fukumori, 2022). Consequently, previous studies have mainly examined the associations between the measures and counseling outcomes. However, it remains unclear how these measures relate to focusers’ subjective evaluations, such as perceived psychological changes and satisfaction.

Other studies have examined which interactions in dialogue relate to counseling outcomes. Based on a categorization of focusers’ post-session written reflections, Lietaer and Neirinck (1986) reported that empathy, questions, and paraphrasing were effective listener utterances. However, their study did not analyze actual utterances in the dialogue. Sachse (1990) reported that listeners’ proposals encouraging more in-depth expressions were related to verbalization progress. However, that study analyzed only two types of listener utterances and thus did not fully capture the diverse

interactions in actual Focusing. In contrast, in this study, we constructed a Focusing dialogue dataset and analyzed the diverse utterance patterns in Focusing using DAs.

3.2 Counseling-Related Dialogue Datasets

Here, counseling dialogue datasets are reviewed to clarify the characteristics of our dataset. In recent years, counseling has attracted significant attention in the field of dialogue systems, and many relevant dialogue datasets have been constructed. The counseling roles considered in existing datasets fall into two categories: emotional support and psychotherapeutic intervention. The former includes ESConv (Liu et al., 2021), AugESC (Zheng et al., 2023), and ESCoT (Zhang et al., 2024), which are datasets designed to provide empathetic responses to clients and offer more positive perspectives. The latter includes Anno-MI (Wu et al., 2022), CACTUS (Lee et al., 2024), HealMe (Xiao et al., 2024), and Koko-roChat (Qi et al., 2025), which are datasets designed to provide support in organizing and addressing speakers’ problems. Although some existing counseling datasets include verbalization of concerns and feelings, they are not intentionally designed to capture the verbalization process.

Among counseling techniques, *attentive listening* is particularly close to Focusing. Attentive listening emphasizes listening to the speaker’s narrative rather than directly addressing problems, and it is central to psychotherapeutic approaches such as PCT (Rogers, 1951) and the narrative approach (White and Epston, 1990). Attentive listening dialogues (listening-oriented dialogues) are defined as dialogues in which one speaker listens attentively to the other (Meguro et al., 2009), and datasets targeting such dialogues have also been constructed (Yoshino et al., 2018; Ito et al., 2022). However, whereas listeners in attentive listening dialogues focus on receiving the speaker’s narrative, listeners in Focusing facilitate the focusers’ verbalization. In this study, to clarify the differences between conventional attentive listening dialogues and Focusing dialogues, we compared the distributions of DAs across these two types of datasets.

4 Data Collection

We collected audio and video recordings of Focusing dialogues conducted by two participants, a



Figure 2: Experimental setup. Listener is on the left and focuser is on the right. Top row shows Zoom screenshots, and bottom row shows the setup from behind, with participants in separate rooms.

listener and a focuser, resulting in a dataset of 50 dialogues. The participants interacted via a video chat tool, and they completed questionnaires before and after the dialogue. In addition, DAs were annotated for each utterance in the transcribed dialogues.

4.1 Experimental Settings

Figure 2 shows the experimental setup. The two participants sat in separate soundproof rooms while conducting a dialogue of approximately 20 minutes via Zoom². Each participant’s audio and video were transmitted through a microphone-equipped webcam mounted on top of their PC, and the other participant’s audio was heard through earphones. This setup was designed to avoid audio interference between the participants during recording. An experimenter monitored the dialogue via Zoom from outside the soundproof rooms and informed the participants of the ending time through Zoom’s chat function. The dialogue was terminated at the scheduled ending time (20 minutes after the dialogue’s start), regardless of the progress of Focusing.

We used Zoom’s built-in recording function to record the audio and video of the dialogues. Since the Zoom recordings captured only the upper body, we additionally recorded each participant’s full body with video cameras placed behind their PCs. Furthermore, for future speech analysis, each participant’s speech was recorded with a lavalier microphone attached to the chest.

²<https://zoom.us/>

A Moonlit Path

I was walking along a path where willow branches swayed in the moonlight. As I moved forward, a small stone struck the tip of my shoe. “Ah, something happened to me on this path once,” I murmured, my thoughts drifting back to the past. In the slightly cooler night air, I simply kept walking. A faint scent in the air stirred an old memory. It was not from the distant past, and yet I had somehow forgotten it until that moment. As I turned my attention to it, an indescribable feeling began to rise within me. On this path, far from the bustle of everyday life, even the quiet seemed vivid. I could hear the wind, the chirping of insects, and the steady rhythm of my footsteps. When I looked up, I saw a small but beautiful full moon hanging in the sky.

Figure 3: Short story used in the Focusing work. Originally written in Japanese and translated into English by the authors.

For the recordings, we prepared original Focusing work designed for focusers who were not familiar with Focusing. Specifically, focusers verbalized their impressions and feelings after listening to a short story (Figure 3). The story was created by one of the authors, who is a certified Focusing trainer of The International Focusing Institute³, to elicit a felt sense through emotionally evocative, open-ended descriptions. This Focusing work has two key advantages: first, it does not restrict the metaphors available to the focuser, allowing flexible vocabulary and expressions; second, it does not directly ask about personal experiences, thus minimizing the disclosure of personal information. The story was presented to the focusers as audio synthesized using Google Cloud Text-to-Speech AI⁴ to keep the voice as neutral as possible.

In total, five listeners participated, each conducting 5–15 dialogues. All listeners were certified Focusing trainers with The International Focusing Institute, and each had at least five years of experience in Focusing. The focusers were 50 participants in their 20s to 60s (22 males and 28 females), each taking part in only one dialogue.

4.2 Questionnaires

In this study, we used the EXP checklist and the FMS-18, described in Section 2.2. The FMS-18 was rated on a 4-point scale. In addition, to quantify participants’ subjective evaluations,

³<https://focusing.org/>

⁴<https://cloud.google.com/text-to-speech>

Item	Listener	Focuser
No. of people	5	50
Total utterances	10,033	7,953
Mean utt. len (chars)	18.8 $\pm$ 24.4	13.0 $\pm$ 15.8
Mean duration per utt. (sec)	2.9 $\pm$ 3.5	2.1 $\pm$ 2.4
Mean utt. per dialogue	200.7 $\pm$ 34.8	159.1 $\pm$ 49.5

Table 1: Dialogue statistics (mean $\pm$ standard deviation).

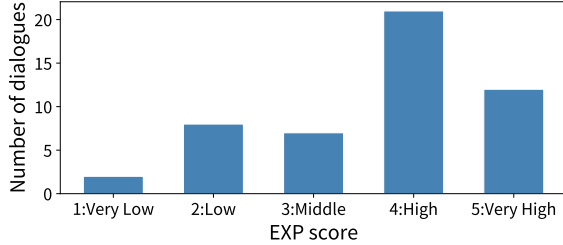


Figure 4: Distribution of EXP scores.

we used the Session Evaluation Questionnaire (SEQ) (Stiles et al., 2002). The SEQ assesses the quality of a counseling session using the Semantic Differential (SD) method. The SEQ consists of two subscales: the *feeling* subscale (10 items) and the *session* subscale (11 items). In this study, we used the Japanese version of the SEQ, the Japanese Session Evaluation Questionnaire (J-SEQ) (Katsuragawa et al., 2013), with all items rated on a 7-point scale.

Participants completed questionnaires before and after the dialogue (pre- and post-dialogue). **Focuser:** Before engaging in the dialogue, the focuser listened to the story’s audio and completed the FMS-18 and the J-SEQ *feeling* subscale. After the dialogue, the focuser completed the full J-SEQ, i.e., both *feeling* and *session* subscales. **Listener:** After the dialogue, the listener completed the EXP checklist and the J-SEQ *session* subscale.

4.3 DA Annotation

First, speech segments were extracted from the Zoom audio using the WebRTC Voice Activity Detector (VAD)⁵. Segments separated by silent intervals of 0.5 seconds or longer were treated as utterances. Each utterance was then automatically transcribed using the openai/whisper-large-v3 model (Radford et al., 2023). Errors in the resulting transcripts were manually corrected by crowd workers while listening to the audio.

For DA annotation, we used the 17 types of DAs proposed for attentive listening dialogue (Meguro

Factor	Scores (mean $\pm$ std.)
<i>attention</i>	16.14 $\pm$ 3.74
<i>acceptance</i>	18.68 $\pm$ 2.61
<i>distance</i>	18.34 $\pm$ 2.86

Table 2: Statistics of FMS-18 for the focusers. Higher scores indicate a stronger Focusing attitude.

et al., 2013); the DA inventory is listed in Table 5. Annotation was performed at the utterance-level by three crowd workers working independently. The inter-annotator agreement (Fleiss’ κ) was 0.65, indicating generally consistent labeling across annotators.

4.4 Statistics of Collected Data

Table 1 shows the statistics of the collected Focusing dialogue dataset. Listeners made more utterances than focusers, and listener utterances were longer. These results suggest that listeners tended to lead the dialogue. Example dialogues with high and low EXP scores are shown in Appendix Tables 11 and 12, respectively.

Figure 4 shows the distribution of EXP scores, with a mean and standard deviation of 3.66 ± 1.14 . The scores were most frequently distributed around 4, with relatively few dialogues scoring below 3. Accordingly, the dataset contains many dialogues in which verbalization progressed to a relatively high level.

Table 2 shows statistics of each FMS-18 factor for the focusers. *Attention* was lower than the other factors, which is consistent with the findings reported by Fukumori (2022). This may be because *attention* is likely to improve through continued Focusing practice, and is therefore less likely to develop through daily life alone. In addition, *attention* had the largest standard deviation, suggesting greater individual differences than the other factors.

Table 3 shows the pre- and post-dialogue statistics of the J-SEQ *feeling* subscale for the focusers. Scores increased for all items from pre- to post-dialogue evaluation, indicating an overall shift in a positive direction. In particular, the large increase in *definite* suggests that focusers themselves perceived their felt sense as becoming clearer through the dialogue. In addition, the relatively large changes in items such as *aroused* and *energetic* suggest increased psychological activation. Table 4 shows post-dialogue statistics of the J-SEQ *session* subscale for the listeners and focusers.

⁵<https://github.com/wiseman/py-webrtcvad>

Item	Pre	Post
<i>sad ↔ happy</i>	4.32±1.19	5.30±1.13
<i>angry ↔ pleased</i>	4.58±0.99	5.12±1.06
<i>calm ↔ excited</i>	1.94±1.08	2.58±1.51
<i>afraid ↔ confident</i>	4.34±1.15	5.06±1.41
<i>unfriendly ↔ friendly</i>	4.74±1.31	5.10±1.45
<i>uncertain ↔ definite</i>	3.70±1.31	4.54±1.62
<i>still ↔ moving</i>	2.68±1.48	3.14±1.57
<i>slow ↔ fast</i>	1.82±1.04	2.60±1.43
<i>peaceful ↔ energetic</i>	2.08±0.97	3.12±1.64
<i>quiet ↔ aroused</i>	1.96±1.23	3.14±1.74

Table 3: Statistics of J-SEQ *feeling* subscale for the focusers (mean ± standard deviation). Higher scores indicate a stronger tendency toward the adjective on the right. Pre and Post indicate ratings before and after the dialogue, respectively. **Bold** font indicates the higher mean for each item.

Item	Listener	Focuser
<i>bad ↔ good</i>	5.28±1.33	6.04±0.90
<i>difficult ↔ easy</i>	4.38±1.43	5.02±1.52
<i>worthless ↔ valuable</i>	5.08±1.51	4.94±1.88
<i>shallow ↔ deep</i>	5.00±1.50	5.68±1.20
<i>tense ↔ relaxed</i>	4.76±1.38	4.28±1.88
<i>unpleasant ↔ pleasant</i>	5.52±0.95	6.10±1.04
<i>empty ↔ full</i>	5.18±1.37	5.16±1.45
<i>weak ↔ powerful</i>	4.86±1.46	5.42±1.18
<i>ordinary ↔ special</i>	5.02±1.27	5.36±1.32
<i>rough ↔ smooth</i>	4.76±1.66	5.30±1.31
<i>uncomfortable ↔ comfortable</i>	5.18±1.22	5.42±1.43

Table 4: Statistics of J-SEQ *session* subscale for the listeners and focusers, following the same notation used in Table 3.

Overall, focusers gave higher ratings than listeners, indicating that they tended to evaluate the dialogue more positively.

Table 5 shows the distribution of DAs for the 16,677 utterances with majority-agreed labels among three annotators, where utterances lacking such labels were excluded. The distribution differed between listeners and focusers: DISC accounted for about half of the focusers’ utterances, while ACK, QUES, CONF, and PARA were more frequent in the listeners’ utterances. These results suggest that focusers primarily verbalized their internal states, whereas listeners mainly supported further verbalization through responsive utterances.

5 Analysis

We analyzed the characteristics of Focusing dialogues using the constructed dataset. First, based on the questionnaires, we examined how measures

Tag	Listener	Focuser
ACKNOWLEDGEMENT (ACK)	42.2%	25.0%
QUESTION (QUES)	18.2%	0.2%
CONFIRMATION (CONF)	6.0%	0.7%
FILLER (FIL)	5.9%	7.4%
PARAPHRASE (PARA)	5.0%	0.0%
PROPOSAL (PROP)	4.3%	0.0%
REPEAT	4.3%	1.0%
INFORMATION (INFO)	3.6%	0.2%
SELF-DISCLOSURE (DISC)	3.4%	52.9%
THANKS	2.6%	1.0%
SYMPATHY (SYM)	1.7%	10.0%
GREETING (GR)	1.2%	0.9%
ADMIRATION (ADM)	0.6%	0.1%
APPROVAL (APR)	0.4%	0.0%
OTHER	0.3%	0.2%
APOLOGY	0.3%	0.2%
NON-SYMPATHY	0.1%	0.2%

Table 5: Distribution of DAs for the listeners and focusers.

of verbalization progress, subjective evaluations, and Focusing attitudes were associated with each another. We then analyzed utterance patterns in Focusing dialogues from the DA annotations. Finally, based on these analyses, we discussed design implications for Focusing dialogue systems.

5.1 Questionnaire Analysis

For correlation analyses, Spearman’s rank correlation (Spearman, 1904) was used, and the Benjamini-Hochberg false discovery rate correction (FDR-BH) (Benjamini and Hochberg, 1995) was applied to control for multiple comparisons⁶.

5.1.1 Relationships Between Verbalization Progress and Subjective Evaluations

Figure 5 shows correlations between EXP score and the J-SEQ *session* subscale for listeners and focusers. For listeners, most items showed significant correlations with the EXP score, suggesting that higher EXP scores were associated with more positive subjective evaluations overall. In contrast, for focusers, significant correlations with EXP score were limited to a few items. These results suggest that, compared with listeners, the association between verbalization progress and subjective evaluation was generally weaker for focusers and, moreover, that they did not necessarily align.

To further examine differences between listener and focuser subjective evaluations, we analyzed

⁶Significance is denoted as * $p < .05$ and ** $p < .01$.

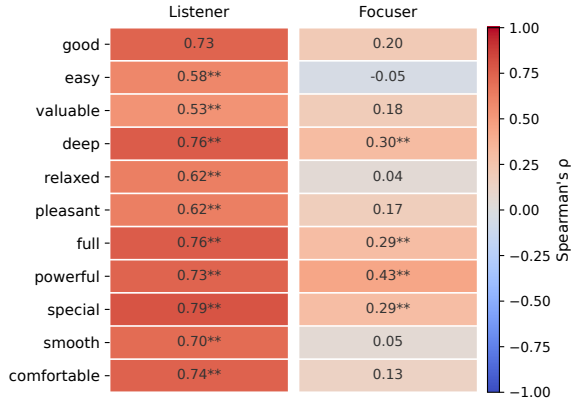


Figure 5: Correlations between EXP score and J-SEQ *session* subscale for the listeners and focusers. The degree of correlation increases from blue to red.

the correlations between their J-SEQ *session* subscale separately for high- and low-EXP groups. Dialogues with an EXP score ≥ 4 were assigned to the high-EXP group, and those with a score ≤ 3 to the low-EXP group. Consequently, the high-EXP group showed significant positive correlations for five items. By contrast, correlations in the low-EXP group were generally weak, with only two items showing significant positive correlations. These results suggest that listener and focuser evaluations are more closely aligned in dialogues with greater verbalization progress. The complete results are presented in Table 13 in Appendix.

5.1.2 Effects of Focusers' Focusing Attitudes on Focusing Outcomes

First, to examine the association between the focusers' Focusing attitude and verbalization progress, we analyzed the correlations between FMS-18 and EXP score. Accordingly, *attention* was moderately correlated with EXP score ($\rho = .45$), whereas *acceptance* ($\rho = -.06$) and *distance* ($\rho = .04$) were not correlated with EXP score. This suggests that higher attention to the felt sense in focusers is associated with greater verbalization progress.

Then, to investigate the association between focusers' Focusing attitudes and speakers' subjective evaluations, we also analyzed the correlations between FMS-18 and the J-SEQ *session* subscale. Here, *attention* showed significant correlations with seven items of the J-SEQ *session* subscale for listeners, and with two items for focusers. These results suggest that higher *attention* in focusers is associated with more positive evaluations

	<i>Focusing dlg.</i>		<i>Listening dlg.</i>	
	Listener	Focuser	Listener	Speaker
DISC	3.4%	52.9%	44.5%	57.3%
INFO	3.6%	0.2%	1.4%	5.2%
ACK	42.2%	25.0%	12.3%	6.9%
QUES	18.2%	0.2%	25.8%	14.0%
SYM	1.7%	10.0%	3.7%	3.3%
GR	1.2%	0.9%	9.8%	9.6%
OTHER	29.5%	10.6%	2.5%	3.7%

Table 6: Distribution of DAs for the listeners and focusers in *focusing* dialogues, and for the listeners and speakers in attentive *listening* dialogues (Meguro et al., 2009).

Tag	High	Low	Sig.
FIL	5.0%	7.0%	**
APR	0.0%	1.0%	*
PROP	5.0%	3.0%	**
ACK	44.0%	38.0%	**
DISC	3.0%	5.0%	**
PARA	6.0%	4.0%	**
QUES	16.0%	22.0%	**

* $p < .05$, ** $p < .01$ (FDR-BH corrected)

Table 7: Distribution of listener DAs in the high- and low-EXP score groups. Only tags with statistically significant differences are shown. Significance was tested using two-proportion z-tests. **Bold** font indicates the higher proportion for each tag.

of the dialogue, particularly from listeners. The complete results are presented in Table 14 in Appendix.

5.2 Linguistic Analysis

Table 6 compares the distribution of DAs in our Focusing dialogues with that in the attentive listening dialogues reported by Meguro et al. (2009)⁷. In both types of dialogue, the focusers or speakers primarily engages in DISC, whereas listener behavior differs substantially: Listeners in Focusing dialogues mainly used ACK and rarely used DISC, while listeners in attentive listening dialogues frequently used DISC. In addition, OTHER was more frequent in Focusing dialogues, and it included many case of CONF, PARA, and PROP. This suggests that listeners in Focusing dialogues support the focuser's verbalization through responsive utterances, whereas listeners in attentive listening dialogues promote it through their own self-disclosure.

Next, we analyzed differences in listener DAs between high- and low-EXP groups. Table 7

⁷We grouped DAs not reported in Meguro et al. (2009) into OTHER.

Speaker	Tag	High	Low	Sig.
Focuser	ACK	23.0%	27.0%	**
	DISC	57.0%	50.0%	**
Listener	ADM	0.0%	1.0%	**
	PROP	5.0%	4.0%	**
	ACK	44.0%	41.0%	*
	CONF	7.0%	6.0%	*
	QUES	16.0%	20.0%	**

* $p < .05$, ** $p < .01$ (FDR-BH corrected)

Table 8: Distribution of DAs for listeners and focusers in the high- and low-FMS-18 score groups, following the same conventions used in Table 7.

shows the distribution of DAs for each group. In the high-EXP group, listeners used significantly more ACK, PARA, and PROP but significantly fewer QUES and DISC. In addition, an analysis of adjacent listener-to-focuser utterance pairs showed that QUES $\rightarrow$ DISC was more frequent in the low-EXP group, whereas PARA $\rightarrow$ SYM and ACK $\rightarrow$ DISC were more frequent in the high-EXP group. These results suggest that, in dialogues with lower verbalization progress, listeners use questions to encourage verbalization, whereas in dialogues with greater progress, they support the focuser’s verbalization through responsive utterances, thus facilitating further verbalization progress.

We also analyzed the differences in DAs between high- and low-FMS groups. Dialogues were divided into these high- and low-FMS groups based on the mean total FMS-18 score (53.16), calculated as the sum of the three factor scores. Table 8 shows the distribution of DAs for each group. For focusers, the high-FMS group showed more DISC, whereas the low-FMS group showed more ACK. For listeners, the high-FMS group showed more PROP and ACK, whereas the low-FMS group showed more QUES. In addition, for adjacent listener-to-focuser utterance pairs, the low-FMS group showed more QUES $\rightarrow$ DISC, while the high-FMS group showed more ACK $\rightarrow$ DISC. These results suggest that, depending on the focuser’s Focusing attitude, listeners vary in how they support verbalization: When the focuser’s Focusing attitude is lower, they use questions more often, whereas when it is higher, they rely more on responsive utterances.

5.3 Implications for System Design

Finally, based on the above analysis results, we discussed design implications for Focusing dia-

logue systems. The correlations between EXP score and the J-SEQ *session* subscale showed that greater verbalization progress does not necessarily lead to more positive evaluations of the dialogue by focusers. Therefore, Focusing dialogue systems should optimize not only verbalization progress but also users’ subjective evaluations, since positive evaluations are important for continued use.

Moreover, given that the distribution of listener DAs differs substantially between Focusing and attentive listening dialogues, dialogue strategies developed for attentive listening dialogue systems cannot be directly applied to Focusing dialogue systems. Focusing dialogue systems should support the user’s verbalization through responsive utterances, rather than promoting it through its own self-disclosure.

Our analysis showed that listener DA usage differed by FMS-18 score and that the *attention* factor was positively correlated with EXP score. Because the user’s Focusing attitude can be assessed in advance through questionnaires, it may be useful for determining the system’s dialogue strategy. Therefore, the system should adapt its dialogue strategy according to the user’s Focusing attitude, especially their *attention*: more questions for users with lower *attention*, and more responsive utterances for those with higher *attention*.

6 Conclusion

In this study, we constructed a multimodal Focusing dialogue dataset toward developing Focusing dialogue systems. The dataset consists of 50 Focusing dialogues between trained listeners and focusers. Each dialogue contains EXP scores, FMS-18 scores, J-SEQ ratings, and annotated DAs. We then used this dataset to analyze how verbalization progress was related to the subjective evaluations of both listeners and focusers, as well as to the focusers’ Focusing attitude. The results showed that focusers’ subjective evaluations were not necessarily correlated with verbalization progress. In contrast, focusers’ Focusing attitudes toward attending to their felt sense were correlated with verbalization progress. Furthermore, analyses based on DAs revealed that both listener and focuser behaviors varied depending on verbalization progress and Focusing attitude. Finally, based on these findings, we discussed design implications for Focusing dialogue systems.

Limitations

This study has several limitations. First, the dataset is limited in scale and coverage. It consists of 50 dialogues based on a single Focusing work, all conducted in Japanese. The DA annotation in the dataset also requires further refinement, as the DA tags were originally proposed for attentive listening and may require customization to better capture the characteristics of Focusing dialogues. Future work should examine the generalizability of the findings using larger and more diverse datasets that include multiple languages, different Focusing works, and DA tags tailored to Focusing dialogues. Second, our analyses relied on verbal information and did not fully capture non-verbal information such as bodily, vocal, and facial expressions. Bodily felt sense is central to Focusing, so such non-verbal cues may provide useful indicators for examining how verbalization progresses. In addition, our DA analysis focused primarily on DA bigrams and did not fully capture longer-range temporal dynamics across the dialogue. Therefore, future analyses should examine such non-verbal cues and broader DA sequences to better understand the characteristics of Focusing dialogues. Third, although our analysis of the collected dialogues provided design implications for Focusing dialogue systems, their feasibility has not yet been validated in an actual dialogue system. In particular, given the rapid progress of large language models (LLMs), it is important to examine how well they can simulate Focusing dialogues. For future work, we plan to develop an LLM-based Focusing dialogue system and evaluate its outputs through comparison with the dialogues collected in this study. Finally, the collected dataset is not yet publicly available and is under preparation for public release.

Ethical Considerations

All experiments were approved by the research ethics committee of our institution. Before participation, all participants were fully informed about the purpose and procedures of the experiments, and they provided voluntary informed consent. In addition, to minimize the disclosure of personal information during the dialogues, we designed the Focusing work to avoid elicitation of detailed personal experiences.

Acknowledgments

This work was supported by JST ACT-X, Japan, Grant Number JPMJAX23CH.

References

- Tsuyoshi Aoki and Akira Ikemi. 2014. The Focusing Manner Scale: its validity, research background and its potential as a measure of embodied experiencing. *Person-Centered & Experiential Psychotherapies*, 13(1):31–46.
- Lisa Feldman Barrett, James J. Gross, Travis C. Christensen, and Michele Benvenuto. 2001. Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition & Emotion*, 15(6):713–724.
- Yoav Benjamini and Yosef Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300.
- Leslie A. Ellis. 2014. Living the dream: The evolution of Focusing-Oriented Dream Work. In *Emerging Practice in Focusing-Oriented Therapy*, pages 166–177. Jessica Kingsley Publishers.
- Hideaki Fukumori. 2022. Examination of the construct validity of the Focusing Manner Scale 18 (FMS-18) with exploratory and confirmatory factor analysis. *Person-Centered and Experiential Psychotherapies*, 21(1):57–72.
- Eugene T. Gendlin. 1961. Experiencing: A variable in the process of therapeutic change. *American Journal of Psychotherapy*, 15:233–245.
- Eugene T. Gendlin. 1981. *Focusing*, 2 edition. Bantam Books.
- Eugene T. Gendlin. 1996. *Focusing-Oriented Psychotherapy: A Manual of the Experiential Method*. The Guilford Press.
- Marta G. Gomes, Eunice R. Silva, and João Salgado. 2024. Focusing with colorectal cancer patients: a pilot study of a brief online group intervention. *Frontiers in Psychology*, 15:1339823.
- Marion N. Hendricks. 2002. Focusing-oriented / experiential psychotherapy: Research and practice. In *Humanistic Psychotherapies: Handbook of Research and Practice*, pages 221–251.
- Akira Ikemi. 2005. Carl Rogers and Eugene Gendlin on the Bodily Felt Sense: What they share and where they differ. *Person-Centered & Experiential Psychotherapies*, 4(1):31–42.

- Akira Ikemi. 2024. The distinction between ‘focusing-alpha’ and ‘focusing-beta’, the practice of ‘Crossing with Animals’ and theoretical considerations on some aspects of Experiencing. In *Proceedings of the 16th World Conference for Person-Centered and Experiential Psychotherapy and Counseling*, pages 30–31.
- Koichiro Ito, Masaki Murata, Tomohiro Ohno, and Shigeki Matsubara. 2022. Construction of responsive utterance corpus for attentive listening response production. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference (LREC)*, pages 7244–7252.
- Taisuke Katsuragawa, Yoshihiko Kunisato, Jun Kanno, and Kazuyoshi Sasaki. 2013. Development of the Japanese version of the session evaluation questionnaire (J-SEQ): Therapist-level analyses. *The Japanese Journal of Personality*, 22(1):73–76. (in Japanese).
- Donald J. Kiesler. 1971. Patient experiencing and successful outcome in individual psychotherapy of schizophrenics and psychoneurotics. *Journal of Consulting and Clinical Psychology*, 37(3):370–385.
- M. H. Klein, P. L. Mathieu, E. T. Gendlin, and D. J. Kiesler. 1969. *The Experiencing Scale: A Research and Training Manual*, volume 1. University of Wisconsin Extension Bureau of Audiovisual Instruction.
- Suyeon Lee, Sunghwan Kim, Minju Kim, Dongjin Kang, Dongil Yang, Harim Kim, Minseok Kang, Dayi Jung, Min Hee Kim, Seungbeen Lee, Kyong-Mee Chung, Youngjae Yu, Dongha Lee, and Jinyoung Yeo. 2024. CACTUS: Towards psychological counseling conversations using cognitive behavioral theory. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 14245–14274.
- Germain Lietaer and Marc Neirinck. 1986. Client and therapist perceptions of helping processes in client-centered/experiential psychotherapy. *Person-Centered Review*, 1(4):436–455.
- Siyang Liu, Chujie Zheng, Orianna Demasi, Sahand Sabour, Yu Li, Zhou Yu, Yong Jiang, and Minlie Huang. 2021. Towards emotional support dialog systems. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP)*, pages 3469–3483.
- Toyomi Meguro, Ryuichiro Higashinaka, Kohji Dohsaka, Yasuhiro Minami, and Hideki Isozaki. 2009. Analysis of listening-oriented dialogue for building listening agents. In *Proceedings of the Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*, pages 124–127.
- Toyomi Meguro, Yasuhiro Minami, Ryuichiro Higashinaka, and Kohji Dohsaka. 2013. Learning to control listening-oriented dialogue using partially observable Markov decision processes. *ACM Transactions on Speech and Language Processing (TSLP)*, 10(4):15.
- Yuko Morikawa, Koji Nagano, and Hideaki Fukumori. 2014. Construction of revised FMS (The Focusing Manner Scale) and confirmation of its reliability and validity. *Journal of the Faculty of International Studies of Culture, Kyushu Sangyo University*, 58:117–135. (In Japanese).
- Rino Ohashi and Akira Ikemi. 2017. The relationship between estimated client experiencing and therapists’ perception of therapy progress. *Psychologist: Kansai University Bulletin of Professional Graduate School of Clinical Psychology*, 7:39–48. (in Japanese).
- Bella Ozomaro, Perel Y. Wein, Sloane Miller, Luz H. Ospina, Marianne Goodman, John Torous, and David Kimhy. 2025. Emotional granularity in health and psychopathology: A scoping review. *Journal of Psychiatric Research*, 190:277–295.
- Antonio Pascual-Leone and Nikita Yeryomenko. 2017. The client “experiencing” scale as a predictor of treatment outcomes: A meta-analysis on psychotherapy process. *Psychotherapy Research*, 27(6):653–665.
- James W. Pennebaker and Cindy K. Chung. 2000. Expressive writing, emotional upheavals, and health. *Psychological Science*, 8(3):162–166.
- Zhiyang Qi, Takumasa Kaneko, Keiko Takamizo, Mariko Ukiyo, and Michimasa Inaba. 2025. Koko-roChat: A Japanese psychological counseling dialogue dataset collected via role-playing by trained counselors. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 12424–12443.
- Alec Radford, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, and Ilya Sutskever. 2023. Robust speech recognition via large-scale weak supervision. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, pages 28492–28518.
- Nathaniel J. Raskin and Carl R. Rogers. 2005. Person-centered therapy. In Raymond J. Corsini and Danny Wedding, editors, *Current Psychotherapies*, 7 edition, pages 130–165. Thomson Brooks/Cole Publishing Co.
- Carl R. Rogers. 1951. *Client-Centered Therapy; Its Current Practice, Implications, and Theory*. Houghton Mifflin.
- Carl R. Rogers. 1961. *On Becoming a Person: A Therapist’s View of Psychotherapy*. Houghton Mifflin.

- Rainer Sachse. 1990. The influence of therapist processing proposals on the explication process of the client. *Person-Centered Review*, 5(3):321–344.
- Charles Spearman. 1904. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101.
- William B. Stiles, Laura E. Gordon, and Julie A. Lani. 2002. Session evaluation and the session evaluation questionnaire. In *Counseling Based on Process Research*, pages 325–343. Allyn & Bacon.
- Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Dan Jurafsky, Paul Taylor, Rachel Martin, Carol Ess-Dykema, and Marie Meteer. 2000. Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics*, 26(3):339–373.
- Michael White and David Epston. 1990. *Narrative Means to Therapeutic Ends*. W.W. Norton & Company.
- Ming Yu Claudia Wong. 2025. The potential psychological advantages of the focusing-oriented therapy: A pilot quasiexperimental study of “proficiency as focusing partner” training workshop on self-compassion. *International Journal of Clinical Practice*, 2025(1):8466900.
- Zixiu Wu, Simone Balloccu, Vivek Kumar, Rim Helaoui, Ehud Reiter, Diego Reforgiato Recupero, and Daniele Riboni. 2022. Anno-Mi: A dataset of expert-annotated counselling dialogues. In *Proceedings of 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6177–6181.
- Mengxi Xiao, Qianqian Xie, Ziyang Kuang, Zhicheng Liu, Kailai Yang, Min Peng, Weiguang Han, and Jimin Huang. 2024. HealMe: Harnessing cognitive reframing in large language models for psychotherapy. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1707–1725.
- Koichiro Yoshino, Hiroki Tanaka, Kyoshiro Sugiyama, Makoto Kondo, and Satoshi Nakamura. 2018. Japanese dialogue corpus of information navigation and attentive listening annotated with extended ISO-24617-2 dialogue act tags. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC)*, pages 2922–2927.
- Kei Yoshiwara and Hironobu Tsuchiya. 2019. Correlations among focusing attitudes, psychological competitive abilities and public self-consciousness in college athletes. *Person-Centered & Experiential Psychotherapies*, 18(1):85–97.
- Tenggan Zhang, Xinjie Zhang, Jinming Zhao, Li Zhou, and Qin Jin. 2024. ESCoT: Towards interpretable emotional support dialogue systems. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 13395–13412.
- Chujie Zheng, Sahand Sabour, Jiaxin Wen, Zheng Zhang, and Minlie Huang. 2023. AugESC: Dialogue augmentation with large language models for emotional support conversation. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 1552–1568.

A Appendix

A.1 EXP Checklist and FMS-18 Items

We present a list of questionnaire items used in this study. Table 9 shows the EXP checklist, and Table 10 shows the items of the FMS-18 scale.

A.2 Example of a Focusing Dialogue From Our Dataset

Table 11 shows an example from a dialogue with a high-EXP score, whereas Table 12 shows an example from a dialogue with a low-EXP score.

A.3 Detailed Results of Questionnaire Correlation Analysis

Table 13 shows the correlations between listeners’ and focusers’ ratings of the J-SEQ *session* subscale. Table 14 shows the correlations between FMS-18 and ratings of the J-SEQ *session* subscale for the listeners and focusers.

Item	High	Low
<i>empty</i> ↔ <i>full</i>	0.49**	−0.37
<i>difficult</i> ↔ <i>easy</i>	0.48**	0.36
<i>ordinary</i> ↔ <i>special</i>	0.43*	−0.03
<i>shallow</i> ↔ <i>deep</i>	0.40*	0.09
<i>weak</i> ↔ <i>powerful</i>	0.35*	−0.25
<i>rough</i> ↔ <i>smooth</i>	0.34	0.52*
<i>bad</i> ↔ <i>good</i>	0.30	0.50*
<i>tense</i> ↔ <i>relaxed</i>	0.24	−0.32
<i>uncomfortable</i> ↔ <i>comfortable</i>	0.22	−0.05
<i>worthless</i> ↔ <i>valuable</i>	0.21	−0.23
<i>unpleasant</i> ↔ <i>pleasant</i>	0.01	0.14

* $p < .05$, ** $p < .01$ (FDR-BH corrected)

Table 13: Spearman correlations for each item between listeners’ and focusers’ ratings of J-SEQ *session* subscale, analyzed separately for high- and low-EXP groups.

No.	Item
1 Very Low	
1	The client focused only on describing events or facts.
2	No emotional expression was observed in the client's speech.
3	Even when asked about their feelings, the client did not express any emotion.
4	The client described their experiences in an intellectualized manner (e.g., "I think," "I considered"), organizing them cognitively rather than emotionally.
2 Low	
5	The client mainly talked about events or situations, but expressed emotions when asked about their feelings.
6	The client clearly expressed basic emotions (e.g., joy, anger, sadness, and pleasure) in relation to events or situations.
7	The client's emotional expressions were limited to reactions to events.
3 Middle	
8	The client reflected on their feelings and used them to describe their state of being.
9	The client expressed not only general reactions to events but also their own unique way of feeling.
10	Through descriptions of situations or events, the client conveyed aspects of themselves.
11	Expressions of a felt sense were observed in the client's speech.
4 High	
12	The client spoke while searching for expressions that matched their feelings.
13	The client attempted to understand their feelings (e.g., "maybe it's this ... or that ...").
14	The client sometimes paused and searched for words while attending to their feelings.
5 Very High	
15	During the session, the client showed moments of sudden realization (e.g., "Ah, I see").
16	The client appeared to make discoveries, showing emotional shifts such as excitement, a raised voice, laughter, tears, or surprise.
17	The client demonstrated successive discoveries or insights.

Table 9: EXP Checklist II ver.1.0 (Ohashi and Ikemi, 2017). In this checklist, "client" refers to the focuser. Originally written in Japanese and translated into English by the authors.

No.	Item
Attention	
1	Even when I cannot grasp my feelings immediately, I try to pay attention to my inner experiences
4	I take time to clarify what I am feeling and what kinds of emotions I may have
7	Every now and then in my daily life, I slowly ask myself, "How am I feeling inside?"
10	I try to sense subtle changes inside myself
13	In my daily life, I hold some time to quietly pay attention to my inner experiences
16	I have a sense that I am living my life being in touch with my inner experiences
Acceptance	
2	I express myself while trusting my feelings
5	I think that I can trust my senses
8	Whatever feelings may come up inside myself, I try to accept them as they are
11	I think it is OK whatever feelings I may have inside
14	I take actions in accordance with my feelings
17	What I express with words are in sync with my feelings
Distance	
3	When I am stuck, I take a pause and stop thinking
6	When I have some concerns in my life, I try to take a distance from them
9	When I have some concerns, I can take a pause and keep a distance
12	When a situation gets stuck, I try to stand back and take a look at the whole situation
15	When some troubles surface in my daily life, I try not to overthink
18	When I encounter a problem, I try to take some time and calm down

Table 10: The Focusing Manner Scale 18 (FMS-18) (Yoshiwara and Tsuchiya, 2019), shown here in the English version reported in Fukumori (2022). The scale's items are rated on a four-point Likert scale.

Speaker	Utterance
Focuser	When I imagine myself walking along a moonlit path, a clear and pleasant feeling comes over me.
Listener	A clear and pleasant feeling. How might you describe what comes beyond the words “pleasant feeling”?
Focuser	It feels transparent, refreshing, and somehow free of impurity.
Listener	Free of impurity, I see.
Focuser	It feels as if the things that had been piling up in my daily life, like snow, are all being cleared away.
Listener	A feeling of becoming lighter.
Focuser	Yes, I feel lighter, and the more I talk about it, the better it feels.

Table 11: Example of a Focusing dialogue from the constructed dataset with a high-EXP score. Fillers and disfluencies are removed for readability. Originally written in Japanese and translated into English by the authors.

Speaker	Utterance
Listener	Various bodily sensations can arise, such as a feeling of heaviness around the stomach.
Focuser	I see. Right now, it does not feel unpleasant; rather, it feels quiet and calm.
Listener	When you bring your attention to that quiet and calm feeling, how does it feel?
Focuser	That is difficult. I cannot put it into words right away.
Listener	Yes, I understand. I think these kinds of feelings often become clearer when we take time and attend to them slowly.
Listener	Would it be possible to bring a little attention to the sense that you have not been able to relax much lately, or to the feeling of nostalgia?
Focuser	I am not sure I understand what you mean.
Listener	I’m sorry. Let me rephrase the question.

Table 12: Example of a Focusing dialogue from the constructed dataset with a low-EXP score, following the same notation used in Table 11.

FMS-18 factor	Item	Listener ρ	Focuser ρ
<i>attention</i>	<i>uncomfortable ↔ comfortable</i>	0.42**	0.40**
<i>attention</i>	<i>rough ↔ smooth</i>	0.40**	
<i>attention</i>	<i>difficult ↔ easy</i>	0.39**	
<i>attention</i>	<i>ordinary ↔ special</i>	0.35*	
<i>attention</i>	<i>bad ↔ good</i>	0.35*	
<i>attention</i>	<i>weak ↔ powerful</i>	0.32*	
<i>attention</i>	<i>empty ↔ full</i>		0.32*
<i>attention</i>	<i>tense ↔ relaxed</i>	0.30*	

* $p < .05$, ** $p < .01$ (FDR-BH corrected)

Table 14: Spearman correlations between FMS-18 and ratings of J-SEQ *session* subscale for the listeners and focusers. Only statistically significant items are shown.