The Effect of Reading Progress Feedback on Working Alliance and Counseling Outcomes

Authors: Sun Qiwu, Wu Caizhi, Yu Lixia, Wang Weixin, Guocheng Shen, Yu Lixia

Date: 2020-12-10T10:13:48+00:00

Abstract

Progress feedback refers to an intervention that provides counselors with standardized measurement results and interpretations about the client, aiming to correct biases in counselors' subjective evaluations of counseling effectiveness and improve counseling outcomes. Given that the standard "face-to-face" feedback procedure between counselor and client does not conform to Chinese sociocultural customs, this study utilized a multilevel structural equation model to analyze the effects of non-"face-to-face" progress feedback in naturalistic settings on the working alliance and counseling effectiveness. The results revealed that at the between-group level, the feedback group exhibited superior working alliance quality; at the within-group level, the working alliance and symptoms were mutually predictive. At termination, the feedback group demonstrated better outcomes in depressive symptoms and perceived helpfulness of counseling. The conclusion is that progress feedback positively influences the working alliance and counseling effectiveness. This study expands the understanding of progress feedback's mechanism of action from the perspective of the working alliance and Chinese relational dynamics, providing evidence for the effectiveness of progress feedback based on practical intervention.

Full Text

The Effects of Progress Feedback on Working Alliance and Treatment Outcomes

SUN Qiwu[12], WU Caizhi[13], YU Lixia[12], WANG Weixin[4], SHEN Guocheng[4]

[1] School of Psychology, Central China Normal University; Key Laboratory of Adolescent Cyberpsychology and Behavior (Central China Normal University), Ministry of Education, Wuhan 430079, China

[2] Key Laboratory of Human Development and Mental Health of Hubei Province, Wuhan 430079, China
[3] Mental Health Education Center for College Students, Central China Normal University, Wuhan 430079, China

Abstract

Progress feedback is an intervention that involves providing therapists with standardized measurement results and interpretations about clients. Its purpose is to correct biases in therapists' subjective evaluations of treatment effectiveness and improve outcomes. Given that the standard "face-to-face" feedback procedure between therapists and clients does not align with Chinese sociocultural norms, this study employed multilevel structural equation modeling to analyze the effects of non-face-to-face progress feedback on working alliance and treatment outcomes in naturalistic settings. Results showed that at the between-person level, the feedback group demonstrated better working alliance quality; at the within-person level, working alliance and symptoms were mutually predictive. At termination, the feedback group showed superior outcomes in depressive symptoms and perceived helpfulness of counseling. The conclusion is that progress feedback has a positive impact on working alliance and treatment outcomes. This study expands understanding of progress feedback mechanisms from the perspectives of working alliance and Chinese relational dynamics, providing practice-based evidence for the effectiveness of progress feedback interventions.

Keywords: progress feedback, working alliance, treatment outcomes, multilevel structural equation modeling

---

Progress feedback refers to the process of conducting continuous, systematic, and standardized measurements of treatment outcomes (typically after each session) and then providing therapists with these results and their interpretations (Wampold, 2015). Howard et al. (1996) conducted the first study on progress feedback, attempting to help therapists timely understand client progress through standardized measurement and assessment. Over the subsequent two decades, interest in progress feedback interventions has gradually increased. However, most research has focused on intervention effects, with insufficient exploration of why progress feedback works and under what conditions it is effective. Given the potential broad applications of progress feedback in counseling (supervision, novice training, improving treatment outcomes, and early identification of deteriorating or non-responding clients), further investigation of its effects, conditions, and pathways is necessary.

1.1 Effects of Progress Feedback Interventions

Research progress since Howard et al. (1996) has primarily manifested in four areas. First, researchers have successfully simplified progress feedback measure-

ment tools, as traditional assessment methods for client change often contained many items, burdening clients and making it difficult to maintain their interest in completing measures after every session. Second, the application scope of progress feedback has expanded from individual counseling to group therapy and family therapy, and from clients with psychological distress to those with severe mental disorders. Third, progress feedback interventions are effective for identifying clients who are not progressing as expected or who are deteriorating, thereby improving outcomes for these clients. Fourth, continuous and specific progress feedback yields small to medium effect sizes for enhancing treatment effectiveness (Boswell et al., 2015).

A recent meta-analysis on feedback interventions using the Partners for Change Outcome Management System (PCOMS) indicated that progress feedback has small to medium effect sizes for less severe clients but no effect for patients with mental disorders (Østergård & Randa, 2020). Another meta-analysis also found that PCOMS feedback interventions had no effect when outcome measures other than PCOMS itself were used (Pejtersen et al., 2020). Some studies have found negative effects of progress feedback for more severe clients (Errázuriz & Zilcha-Mano, 2018), explaining that continuous negative feedback may undermine confidence in both therapists and clients, leading to negative outcomes. Indeed, most studies supporting the positive effects of progress feedback have been conducted with relatively less severe populations (e.g., university counseling centers) (Østergård et al., 2020). Additionally, specific implementation conditions affect the utility of progress feedback, including adherence to procedures, feedback content and components, the type of assistance provided to non-progressing clients, and certain cultural contextual factors (Lambert & Harmon, 2018).

1.2 Cognitive Mechanisms of Progress Feedback Interventions

Researchers have proposed mechanisms of progress feedback from a cognitive processing perspective. First, when feedback information is inconsistent with therapists' self-evaluation, it creates cognitive dissonance, motivating therapists to reduce this inconsistency (Riemer & Bickman, 2011). Second, like ordinary people, therapists are prone to the "fundamental attribution error"—attributing successful progress to their own abilities while attributing lack of progress to external factors (e.g., the client being "more severe"). Therapists also tend to overestimate treatment effectiveness. Feedback information can correct these blind spots, thereby improving therapeutic efficacy (Macdonald & Mellor-Clark, 2015). Consequently, therapists' attitudes toward progress feedback are important factors influencing its effectiveness (de Jong et al., 2012), as only when therapists hold positive attitudes will they value progress feedback information and potentially overcome biases in their self-assessment of therapeutic progress.

1.3 Working Alliance Theory Perspective

The working alliance refers to a collaborative relationship established between client and therapist, forming a partnership to address the client's difficulties (symptoms), understand their inner experiences (personality structure and dynamics), and modify pathological psychological structures or behaviors. Working alliance components include agreement on therapeutic "goals-tasks" and the emotional bond formed between client and therapist during this process (Bordin, 1994). The working alliance is a core concept in process-outcome research, with meta-analyses demonstrating a stable positive relationship between working alliance and treatment outcomes (Flückiger et al., 2018). Building on this foundation, researchers have attempted to answer two important questions about the alliance-outcome relationship: the causal relationship between working alliance and treatment outcomes, and important moderating variables in this relationship.

Regarding the first question, some researchers view the working alliance as part of treatment outcomes—a byproduct of effective therapy (Barber, 2009; DeRubeis et al., 2005). By repeatedly measuring (typically after each session) symptoms and working alliance, the temporal sequence between alliance changes and symptom changes can be established. Increasing evidence suggests that a good working alliance predicts symptom change rather than being a byproduct of symptom relief (Zilcha-Mano, 2017). Regarding moderating variables, researchers have found that client symptom severity (Errázuriz & Zilcha-Mano, 2018; Zilcha-Mano & Errázuriz, 2015), clients' pre-treatment relationship characteristics (Zack et al., 2015), and client-therapist ratios (Del Re et al., 2012) are important moderators.

Previous progress feedback studies often treated working alliance as an outcome variable, providing feedback to therapists alongside symptom assessments. Recently, researchers have begun exploring the role of progress feedback in the alliance-outcome relationship. Progress feedback information can help therapists adjust their clinical work, identify goal-task incongruences between clients and therapists, and promote a collaborative working relationship (Moltu et al., 2018). In this sense, progress feedback can improve working alliance quality, repair alliance ruptures, and thereby influence the alliance-outcome relationship (Eubanks et al., 2019; Miller et al., 2005). Two recent studies attempted to examine this role (Brattland et al., 2019; Sun et al., 2020). Brattland et al. measured working alliance at mid-treatment and found it mediated the effect of progress feedback. However, they did not collect data after each session, making it impossible to rule out whether symptoms had already changed when measuring working alliance. Sun's study attempted to address this limitation using session-by-session data, but found no differences between feedback and control groups in the alliance-outcome relationship due to small sample size.

1.4 Chinese Relational Perspective

Evidence suggests that Chinese ways of handling interpersonal relationships (relational schemas) influence both working alliance and progress feedback. Influenced by traditional Confucianism, Chinese people tend to be filial to parents, respect elders, and suppress personal needs to maintain family harmony. These schemas transfer to social life as respect for authority (superiors), maintenance of harmonious colleague relationships, and emphasis on "face" (Chen & Chen, 2004). In such contexts, Chinese people prefer indirect communication when expressing disagreement. In counseling situations, therapists represent authority; in university settings, clients often address therapists as teachers, while social norms expect students to show respect for authority and teachers (Hwang, 2000). Therefore, in Chinese counseling contexts, therapists must work harder to identify clients' unexpressed needs, desires, and disagreements that are withheld out of respect (Kuo et al., 2011).

The tendency to view therapists as authorities with greater knowledge and experience also leads Chinese clients to have different expectations than Western clients. For example, Chinese clients expect that one task of the working alliance is to follow the therapist's guidance about their problems (Zhu & Jiang, 2011a). Chinese therapists view guidance as beneficial for clients, and if clients accept this guidance, it may enhance the working alliance and affect treatment outcomes (Duan et al., 2015). However, this may influence clients' engagement in therapy and their sense of responsibility (Zhu & Jiang, 2011a). Currently, Chinese psychological counseling practice is built upon Western theories, yet characteristics of "good clients" such as being expressive, proactive, and open to different ideas reflect Western expectations. This inconsistency between therapists' and clients' expectations about how therapy should proceed represents a culturally imprinted pre-therapy working alliance rupture (Sun et al., 2020).

The first progress feedback effectiveness study in China found that Chinese therapists were unwilling to follow Western standard procedures of conducting progress feedback measurements and discussions "face-to-face" with clients during sessions. They preferred research assistants to administer progress feedback measures and inform them of the results (She et al., 2018). Modified progress feedback can serve as a corrective tool for client-therapist incongruences (alliance ruptures) without requiring face-to-face discussion, thereby maintaining both parties' "face" (dignity) and preserving a workable harmonious relationship. Therapists tend to report feedback information to their supervisors to seek guidance rather than discuss differences or conflicts with clients—an indirect communication style. Under these circumstances, a reasonable hypothesis is that progress feedback will improve the working alliance between Chinese clients and therapists, thereby affecting treatment outcomes.

1.5 Research Hypotheses

Based on the above discussion, this study further examines the effects, conditions, and pathways of progress feedback. Specifically, it addresses three questions: (1) What are the effects of using progress feedback in routine practice at Chinese university counseling centers? We hypothesize that progress feedback used in routine practice at university counseling centers has positive effects. (2) How do therapists' behaviors and attitudes toward using progress feedback information influence its effectiveness? We hypothesize that both therapists reading progress feedback information and their positive attitudes toward progress feedback have positive effects on intervention outcomes. (3) What are the pathways through which progress feedback works? We hypothesize that progress feedback improves the working alliance between Chinese clients and therapists, thereby affecting treatment outcomes.

2.1 Research Procedure

This project collected data at the Mental Health Education Center for College Students of Central China Normal University (hereinafter "the Counseling Center"). The project was approved by the university's ethics committee. Data collection occurred from September 2018 to January 2019. To serve more clients with counseling needs, the Counseling Center stipulates a maximum of six counseling sessions per semester. For clients deemed by therapists to require additional sessions, the Center would increase their session allotment. Clients met with therapists once weekly for 50-minute sessions at no cost.

The Counseling Center operates on an appointment system. Clients first complete an appointment registration form, which includes basic information, the Patient Health Questionnaire (PHQ-9) (Kroenke et al., 2001; Bian et al., 2009), the Generalized Anxiety Disorder scale (GAD-7) (Spitzer et al., 2006; He et al., 2010), and informed consent for participation. After registration, the Center arranges a triage assessment by a psychiatrist from an affiliated hospital within two weeks, who provides treatment recommendations or counseling referrals. Clients complete the Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM) (Evans et al., 2002) before each session and the Working Alliance Questionnaire (WAQ) (Zhu & Jiang, 2011b) after each session. At treatment termination, clients complete PHQ-9, GAD-7, CORE-OM, and a post-treatment survey.

Before the study, the Counseling Center provided therapists with one hour of training covering the function of progress feedback, measurement tools used, and how to interpret feedback information. In a pilot survey, approximately 80% of therapists reported that progress feedback would create considerable pressure. To avoid excessive pressure, the Center encouraged but did not require therapists to use progress feedback information. Research assistants emailed assessment results to therapists on the day of each session, including CORE-OM and WAQ scores, score interpretations, CORE-OM change trajectories,

whether changes represented reliable change indices, and graphical curves of score changes. After data collection, therapists completed a post-study survey about progress feedback.

Inclusion criteria for clients were: (1) clients deemed suitable for counseling after triage, excluding those requiring medication or hospitalization; (2) clients with two or more sessions, excluding those with only one session; and (3) clients who saw the same therapist, excluding those who saw different therapists sequentially. Of 524 clients who scheduled appointments, 61 were not assigned to therapists (e.g., referred for medical treatment, scheduling conflicts), 18 refused participation, and 91 had only one session. Among clients with two or more sessions, four saw two different therapists. Based on post-study survey responses about whether therapists read progress feedback information, feedback and control groups were determined (Figure 1 [Figure 1: see original paper]).

2.2 Participants

Therapists: Forty-eight therapists participated, including 11 males and 37 females. Fourteen were master's-level counseling interns, eight were full-time Counseling Center staff, and 26 were part-time contracted therapists. Therapist ages ranged from 24 to 54 years ($M = 34.00$, $SD = 6.70$), with clinical experience ranging from 0.50 to 20 years ($M = 5.50$, $SD = 3.50$). Using 5-point Likert scales (1 = not at all; 5 = completely), therapists self-rated their orientation across four theoretical approaches: humanistic/existential ($M = 3.00$, $SD = 0.50$), cognitive-behavioral ($M = 2.25$, $SD = 0.40$), psychodynamic ($M = 3.23$, $SD = 0.86$), and integrative ($M = 3.72$, $SD = 0.53$).

Clients: The final analyzed sample included 350 clients (see Figure 1). Among them, 255 (72.90%) were female and 95 were male. Client ages ranged from 18 to 33 years ($M = 21.82$, $SD = 2.53$). Presenting problems included academic issues, interpersonal problems, emotional difficulties, and family problems. Approximately 11.60% of clients reported previous psychiatric treatment.

2.3 Measures

Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM): The CORE-OM contains 34 items assessing well-being, symptoms, social functioning, and risk. Clients rate their experiences over the past week on a 0-4 Likert scale (0 = not at all, 1 = only occasionally, 2 = sometimes, 3 = often, 4 = always). The CORE-OM has multiple versions. This study used the 34-item version before the first session and the 10-item short version for subsequent progress feedback assessments. The Chinese version translated by Zhang et al. (2020) was used. In this sample, the correlation between pre-treatment CORE-OM and PHQ-9 was $r = 0.65$ ($p < 0.001$), and with GAD-7 was $r = 0.59$ ($p < 0.001$), supporting concurrent validity. Internal consistency reliability for pre-treatment CORE-OM was 0.93. Total CORE-OM scores were used in analyses.

Working Alliance Questionnaire (WAQ): The WAQ includes 12 items assessing three factors: engagement, goal-task agreement, and emotional bond. Developed by Zhu and Jiang (2011b) based on qualitative research with Chinese clients, clients rate each item on a 1-5 scale (1 = rarely, 2 = sometimes, 3 = often, 4 = most of the time, 5 = always) based on the session. Internal consistency reliability for WAQ measured after the first session was 0.87. Total WAQ scores were used in analyses.

Patient Health Questionnaire (PHQ-9): The PHQ-9 contains nine items assessing depressive symptoms over the past two weeks on a 0-3 scale (0 = not at all, 1 = several days, 2 = more than half the days, 3 = nearly every day). Internal consistency reliability was 0.88. Total PHQ-9 scores were used.

Generalized Anxiety Disorder Questionnaire (GAD-7): The GAD-7 contains seven items assessing anxiety symptoms over the past two weeks on a 0-3 scale (0 = not at all, 1 = several days, 2 = more than half the days, 3 = nearly every day). Internal consistency reliability was 0.92. Total GAD-7 scores were used.

Post-treatment Survey: The therapist post-survey asked, "When do you usually read progress feedback information?" with options: A) usually no time, B) find time on the day of session, C) before each session, D) did not receive feedback. Therapists selecting A or D were assigned to the control group. Therapists' attitudes toward feedback usefulness were measured by "How useful do you think the feedback information is?" (0 = not at all useful, 9 = extremely useful) (M = 5.19, SD = 1.70). The client post-survey included: "How helpful do you think the counseling was?" (1 = not at all helpful, 5 = extremely helpful) (N = 285, M = 3.54, SD = 1.40); "To what extent did you achieve your counseling goals?" (0 = did not achieve goals at all, 10 = completely achieved goals) (N = 285, M = 7.55, SD = 1.90); and "Your satisfaction with this counseling" (1 = completely dissatisfied, 5 = completely satisfied) (N = 285, M = 4.49, SD = 0.74).

2.4 Analytical Strategy

This study used multilevel linear models to analyze data, decomposing within-person and between-person effects. Multilevel models better handle unbalanced designs with varying numbers of repeated measurements across clients. Within-person effects refer to changes in symptoms measured before each session and working alliance measured after each session for a given client, while between-person effects refer to differences in symptoms and working alliance across different clients. Whether therapists read progress feedback information was a between-person variable.

Multilevel models can estimate within-person and between-person effects through various methods. A common approach is within-person centering—subtracting the within-person mean to eliminate between-person slope influences, such as:

$$CORE_{i,t} = \beta_{0i} + \beta_{1i}(WAI_{i,t-1} - \overline{WAI}_i) + \varepsilon_{i,t} \quad (1)$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01}\overline{WAI}_i + u_{0i} \quad (2)$$

where β_{0i} is the random intercept and $\overline{WAI}_i$ represents between-person differences. With small numbers of repeated measurements (as in this study), this centering method can yield biased estimates of $\overline{WAI}_i$. More accurate estimates can be obtained by treating within-person variable means as latent variables u_{xi} (Asparouhov & Muthén, 2019). The multilevel model can then be written as:

$$CORE_{i,t} = \beta_{0i} + \beta_{1i}(WAI_{i,t-1} - \overline{u}_{xi}) + \varepsilon_{i,t} \quad (3)$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01}\overline{u}_{xi} + u_{0i} \quad (4)$$

Allowing the coefficient of $WAI_{i,t-1}$ predicting $CORE_{i,t}$ (here β_{1i}) to vary across individuals creates a mixed model with both fixed and random effects:

$$\beta_{1i} = \gamma_{10} + u_{1i} \quad (5)$$

For time-series data, autoregressive issues must also be considered, as shown in equations 9-10 (here r_y and r_x) (Asparouhov & Muthén, 2019). This model is sometimes called an AR(1) model:

$$WAI_{i,t} = WAI_{b,i} + WAI_{w,it} \quad (6)$$

$$CORE_{i,t} = \beta_{0i} + \beta_{1i}WAI_{w,it} + \varepsilon_{i,t} \quad (7)$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01}WAI_{b,i} + \varepsilon_i \quad (8)$$

$$\varepsilon_{i,t} = r_y\varepsilon_{i,t-1} + \delta_{it} \quad (9)$$

$$WAI_{w,it} = r_{xWAI_{w,i,t-1}} + \xi_{it} \quad (10)$$

Bayesian methods can be used to obtain standardized solutions and confidence intervals for within-person and between-person effects. Bayesian estimation employs Markov chain Monte Carlo (MCMC) methods. The potential scale

reduction (PSR) statistic assesses model convergence—when PSR < 1.05, iteration terminates as continued computation would yield equivalent parameter estimates. The Deviance Information Criterion (DIC) is used for model comparison; smaller DIC values indicate better models. In Bayesian estimation, posterior predictive p-values (PPP) indicate model fit, representing similarity between two MCMC chains, with non-significant p-values indicating acceptable model-data fit. For multilevel models with between-person slopes, Mplus (Muthén & Muthén, 2017) does not report PPP but provides DIC and pD (estimated number of parameters) as model-data fit indices (Zyphur & Oswald, 2013).

Based on these models, we examined the effect of therapists reading progress feedback information. The primary purpose was to investigate the impact of therapists reading progress feedback, for which we established a multilevel model of progress feedback effects on working alliance and treatment outcomes (Figure 2 [Figure 2: see original paper]).

Missing Data Analysis: Multilevel models analyzing repeated measures with unbalanced designs assume data are missing at random (MAR). This study was conducted in naturalistic settings, making it impossible to determine whether missing data resulted from clients achieving their goals and terminating early or from dropout. Non-random missingness can be examined using pattern mixture models (Little, 1993) and selection models (Wu & Carroll, 1988). Given our focus on how therapists reading progress feedback affects working alliance and symptoms, we used pattern mixture models (Little, 1993) for missing data analysis. The basic approach involves coding missing data patterns for working alliance and symptoms as dummy variables, then including these as covariates in the multilevel model to examine their effects on between-person latent variables for working alliance and symptoms.

Statistical Power Analysis: Monte Carlo simulation methods were used to calculate statistical power, allowing parameter manipulation to study model specification effects. Based on this study's sample characteristics, we set observation counts for 2-6 sessions at 50, 60, 70, 80, and 100 respectively, using actual estimates from data with two or more sessions to set model parameters, with 100 replications. Monte Carlo simulation uses "true" parameter values to calculate "estimated" values across replications, enabling calculation of relative bias between true and estimated values and statistical power for each parameter. Results showed that at the between-person level, when therapists reading progress feedback had a medium effect on working alliance (standardized coefficient = 0.38), statistical power was 98% with 0.80% relative bias from the true coefficient.

All models were analyzed using Mplus 8.1 (Muthén & Muthén, 2017); other statistics used Stata 15 (StataCorp, 2017).

3.1 Descriptive Statistics and Preliminary Analysis

Descriptive statistics for working alliance and symptoms are presented in Table 1 . The intraclass correlation coefficient (ICC) for CORE-OM was 0.64, indicating that 64% of symptom variance was between-person and 36% within-person. The ICC for WAQ was 0.69, indicating that 69% of working alliance variance was between-person and 31% within-person.

3.2 Effects of Therapists Reading Progress Feedback on Working Alliance, Treatment Outcomes, and Their Relationship

We established the multilevel linear model shown in Figure 2. Model 1 included within-person autoregressive effects for symptoms and working alliance (S1 and S2) (DIC = 11194.02, pD = 2252.18). Model 2 added feedback effects on these slopes (DIC = 11173.29, pD = 2211.49). Table 2 presents the results. At the between-person level, therapists reading progress feedback improved working alliance quality with a medium effect size. At the within-person level, changes in working alliance and symptoms were mutually predictive, with small effect sizes. At the between-person level, therapists reading progress feedback had no significant effects on autoregressive slopes for working alliance or symptoms, on symptoms predicting working alliance, or on working alliance predicting pre-session symptoms in the next session. Working alliance did not significantly predict symptom levels.

3.3 Effects of Therapists' Attitudes Toward Progress Feedback on Working Alliance, Treatment Outcomes, and Their Relationship

Replacing "reading progress feedback" with "feedback usefulness" in Model 2 (DIC = 11196.32, pD = 2238.96) showed that at the between-person level, therapists' usefulness ratings had no significant effects on working alliance, symptom levels, autoregressive slopes for working alliance or symptoms, symptoms predicting working alliance, or working alliance predicting pre-session symptoms in the next session.

3.4 Sensitivity Analysis

Trends in Time-Series Data: We conducted several analyses building on Model 2. First, we included session number as a covariate to control for its effects. When session number was included as a between-person covariate correlated with symptoms, results (DIC = 18370.79, pD = 2469.45) showed significant correlation between session number and symptoms (b = 0.33, posterior SD = 0.116, one-tailed p = 0.003, 95% CI [0.11, 0.56]). The between-person effect of reading progress feedback on working alliance level slightly increased (b = 0.41, posterior SD = 0.183, one-tailed p = 0.003, 95% CI [0.12, 0.85]).

When session number was correlated with working alliance, results (DIC = 18354.41, pD = 2470.77) showed no significant correlation (b = 0.27, posterior

SD = 0.28, one-tailed p = 0.130, 95% CI [-0.47, 0.79]). The between-person effect of reading progress feedback on working alliance level slightly increased (b = 0.46, posterior SD = 0.20, one-tailed p = 0.003, 95% CI [0.05, 0.29]).

Controlling for time-invariant confounds and removing trends in time-series data can yield more reliable results, though this may also remove covariation effects between variables over time (e.g., within-person reciprocal effects between alliance and symptom changes) (Falkenström et al., 2017). Allowing correlation between S1 and S2 (representing covariation between autoregressive slopes for symptoms and working alliance) showed (DIC = 11132.76, pD = 2335.32) no significant correlation (b = -0.27, posterior SD = 0.32, one-tailed p = 0.198, 95% CI [-0.87, 0.31]). After controlling for this effect, the between-person effect of reading progress feedback on working alliance level slightly decreased (b = 0.35, posterior SD = 0.15, one-tailed p = 0.005, 95% CI [0.08, 0.72]). Within-person and between-person results remained similar to Table 2, with reciprocal prediction between working alliance and symptom changes.

Missing Data Analysis: Using all data (2-17 sessions) for missing data analysis would generate up to 2^n-1 dummy variables (e.g., 2^{17}-1 for 17 time points), making model fitting impossible. Given the Center's six-session limit, we used data from 2-6 sessions for missing data analysis. Working alliance had five missing data patterns and symptoms had six, yielding four dummy variables for alliance patterns and five for symptom patterns. Results (DIC = -43388.34, pD = 13594.67) showed that symptom missing data patterns had no effect on between-person symptom levels, and working alliance missing data patterns had no effect on between-person working alliance levels, indicating no significant non-random missingness effects. Other main results are presented in Table 2.

3.5 Post-hoc Analysis

Initial Symptom and Working Alliance Levels Between Groups: Since therapist reading of progress feedback was determined by post-study survey rather than random assignment, we examined initial differences. Results showed no differences in symptom levels (CORE-OM) (Mdiff = 0.72, se = 0.73, 95% CI [-0.73, 2.17], t = 0.98, df = 348) or working alliance quality (Mdiff = -1.07, se = 0.99, 95% CI [-3.04, 0.89], t = -1.08, df = 348). Among 257 clients who completed PHQ-9 and GAD-7 at intake, there were no significant differences in initial depressive symptoms (Mdiff = 0.94, se = 0.94, 95% CI [-0.94, 2.82], t = 0.99, df = 255) or anxiety symptoms (Mdiff = 0.24, se = 0.83, 95% CI [-1.41, 1.89], t = 0.29, df = 255) between groups.

Session Numbers Between Groups: The feedback group improved between-person working alliance quality, which is positively related to symptom reduction. Therefore, we hypothesized that the feedback group would achieve goals more quickly, evidenced by fewer average sessions. Results showed the feedback group had 1,084 session units (M = 3.58, se = 0.07, SD = 2.45) versus 408 ses-

sion units for the control group (M = 4.43, se = 0.18, SD = 3.54), a significant difference (Mdiff = 0.86, se = 0.19, 95% CI [0.49, 1.23], t = 4.53, df = 1490).

Distribution of Intern and Professional Therapists: If the feedback group had more experienced therapists with better counseling efficacy or greater ability to utilize feedback, therapist experience would confound the results. We examined therapist distribution: among 87 control group clients, three saw intern therapists and 84 saw professional therapists (96.50%); among 263 feedback group clients, 98 saw intern therapists and 165 saw professional therapists (62.70%). The control group had significantly more clients with professional therapists than intern therapists (^{2} = 36.41, df = 1, p < 0.001).

Intern vs. Professional Therapists in the Feedback Group: To further examine experience effects, we used the Figure 2 model with therapist type instead of reading feedback on feedback group data. Results (DIC = 8126.68, pD = 1623.39) showed no significant differences at the between-person level between intern and professional therapists in working alliance level (b = 0.21, posterior SD = 0.26, one-tailed p = 0.165, 95% CI [-0.39, 0.75]) or client symptoms (b = 0.16, posterior SD = 0.26, one-tailed p = 0.155, 95% CI [-0.47, 0.45]).

Treatment Termination Outcomes: Using propensity score matching (predictors: pre-treatment CORE-OM-34, PHQ-9, and GAD-7), we examined feedback effects on outcomes. The feedback group (N = 170) had lower PHQ-9 scores than the control group (N = 57) (b = -2.29, se = 1.05, z = -2.18, p = 0.029, 95% CI [-4.36, -0.23]), but effects on GAD-7 and CORE-OM-34 were not significant. The feedback group rated counseling as more helpful (Mdiff = -0.52, se = 0.22, 95% CI [-0.97, -0.08], t = -2.36, df = 225), but groups did not differ significantly on goal achievement or satisfaction.

This study adopted a culturally adapted progress feedback procedure based on Chinese counseling center realities and therapists' attitudes toward routine implementation. Results showed that compared to the control group, therapists reading feedback information had a medium effect size on improving working alliance quality. At the within-person level, working alliance changes and symptom changes were reciprocally predictive with small effect sizes.

Therapists received feedback including symptom and working alliance scores after each session, graphical curves of score changes across sessions, score interpretations, and reliable change indices. These components align with international progress feedback practices (Lambert & Harmon, 2018), but differ in implementation procedures. If feedback procedures do not meet therapists' needs, they can create substantial pressure, increase workload, trigger resistance, and fail to achieve collaborative goals (Moltu et al., 2018; Oanes et al., 2015). By adjusting procedures according to Chinese sociocultural norms based on preliminary surveys, we created "room" and "buffer space" for both parties to think about and resolve inconsistencies, making progress feedback more acceptable and effective.

Researchers recognize that providing feedback information alone is insufficient to improve outcomes and that more specific feedback and problem-solving tools

are needed (Lambert & Harmon, 2018; Lutz et al., 2015). This recognition stems from the U.S. system where therapists, after licensure, must complete continuing education but have no regular supervision requirements. In the Counseling Center where this study was conducted, all therapists receive regular individual or group supervision. Post-study surveys indicated approximately 70% of therapists reported discussing progress feedback with supervisors. Thus, stable and reliable supervision help may partially explain feedback group effects when feedback indicates lack of progress. Only three control group clients were with intern therapists, suggesting interns were more eager for client feedback and used it to improve therapy in supervision. This factor has been overlooked in previous progress feedback research. Future studies should examine intern therapists separately and explore supervision's specific impact on using feedback information.

Therapists reading progress feedback had medium effects on improving working alliance quality, consistent with previous research (Burlingame et al., 2018; McClintock et al., 2017; She et al., 2018). Notably, the control group consisted almost entirely of professional therapists, suggesting some professionals were confident in their subjective assessments and did not find feedback helpful. This aligns with past research showing that 80% of therapists (often independent practitioners) believe their skills rank in the top 20% (Walfish et al., 2012) and that therapists' subjective outcome assessments are unreliable, overestimating success and underestimating failure risk (Boswell & Constantino, 2015; Boswell et al., 2015). These results highlight progress feedback's value even for experienced therapists. Three theoretical explanations account for these effects. First, feedback information creates "alertness" or cognitive dissonance, prompting therapists to re-evaluate and adjust therapy (Riemer & Bickman, 2011). Second, from an alliance perspective, feedback helps therapists identify alliance ruptures—goal-task incongruences and emotional tensions (Eubanks et al., 2019)—enabling identification and resolution. Third, given Chinese indirect communication styles, culturally adapted feedback procedures can help therapists identify unexpressed client wishes and needs or discover goal-task questions withheld out of "respect," making it particularly useful for maintaining harmonious relationships and building or repairing alliance (Kuo et al., 2011; Sun et al., 2020).

At the between-person level, therapists reading progress feedback had no significant effect on symptom levels assessed by CORE-OM-10 but reduced session numbers, consistent with two previous studies (Janse et al., 2017; McClintock et al., 2017). Since feedback and outcome assessment used the same tool, some researchers suggest feedback effects may result from repeated measurement or allegiance effects (Østergård et al., 2020; Pejtersen et al., 2020). Therefore, using measures other than feedback tools provides more valuable information. This study found that the feedback group showed better depressive symptom improvement and higher self-rated counseling helpfulness, indicating positive effects on specific measures—contrasting with a recent meta-analysis (Pejtersen et al., 2020). This suggests future progress feedback research should use out-

come measures different from feedback assessment tools to overcome allegiance effects.

At the within-person level, reciprocal prediction between working alliance and symptom changes aligns with previous research (Falkenström et al., 2017). For the same client, working alliance changes predict subsequent symptom changes, while symptom changes predict post-session working alliance ratings. This indicates the within-person reciprocal alliance-outcome model has cross-cultural stability.

Progress feedback interventions require greater attention to therapist variables. This study found therapists' attitudes did not affect between-person working alliance quality, possibly because single-item measurement did not capture true attitudes—attitude-behavior consistency was low in this sample ($r = 0.21$, $p < 0.001$). To enhance progress feedback effects, intervening on therapist attitudes is a viable approach (de Jong et al., 2012). This suggests substantial room remains to improve therapists' positive attitudes toward feedback and to facilitate easier access to feedback information, potentially increasing effect sizes.

This study did not find mediating effects of working alliance at the between-person level, nor effects of reading feedback on autoregressive slopes for working alliance (S1) or symptoms (S2), on working alliance predicting next-session symptoms (SXY), or on pre-session symptoms predicting post-session working alliance (SYX). Perhaps some unrecognized between-person and within-person interaction exists, or other potential moderators have larger effect sizes than progress feedback (e.g., clients' relationship formation characteristics) (Zack et al., 2015), preventing detection of feedback effects. Future research should explore this further.

This study has good ecological validity, 主要体现在两个方面: first, adapting progress feedback procedures according to Chinese cultural norms regarding relationships and "feedback"; second, conducting the study in naturalistic settings with group assignment based on post-study survey rather than randomization. This design also avoids experimenter effects, as researchers and clients were unaware beforehand which therapists would not read feedback.

This study used longitudinal data, which establishes temporal relationships between variable changes and thus better enables examination of causal relationships. The statistical methods used overcame common biases in longitudinal data analysis and ruled out alternative explanations. Statistical power and missing data analyses excluded threats to validity, making conclusions more credible. Post-hoc analyses ruled out potential confounds and used measures other than feedback tools to assess outcomes—additional strengths.

Study limitations include: first, data collection was limited to one counseling center, affecting external validity; second, clients had relatively mild psychological distress, requiring further research to determine if conclusions hold for more severe cases. Additionally, Chinese cultural factors and social norms were only theoretically discussed, not directly tested. Future research should directly

examine these hypotheses. Furthermore, while symptom change information in feedback (e.g., reliable change, clinically significant change) is relatively understandable, guidance on working alliance changes is less clear. Future studies should consider how to better feedback working alliance information, such as examining alliance development trends, alliance ruptures, and culturally specific manifestations of alliance ruptures in Chinese relational schemas.

This study found that therapists reading progress feedback information improved working alliance quality, enhanced treatment outcomes, and reduced session numbers. Conducted in naturalistic settings, it provides practice-based evidence for intervention effectiveness, suggesting important applied value for progress feedback in routine outcome monitoring at university counseling centers.

References

References are preserved exactly as provided in the original text, including both English and Chinese entries with their respective translations.

Note: Figure translations are in progress. See original paper for figures.