Published in final edited form as:

Psychother Res. 2015 November ; 25(6): 678–693. doi:10.1080/10503307.2015.1059966.

Feedback mechanisms of change: How problem alerts reported by youth clients and their caregivers impact clinician-reported session content

Susan R. Douglas¹, Bae Jonghyuk¹, Ana Regina Vides de Andrade¹, M. Michele Tomlinson², Ryan Pamela Hargraves³, and Leonard Bickman¹

¹Vanderbilt University

²Hope Network

³Emma

Abstract

Objective—This study explored how clinician-reported content addressed in treatment sessions was predicted by clinician feedback group and multi-informant cumulative problem alerts that appeared in computerized feedback reports for 299 clients aged 11 to 18 years receiving home-based community mental health treatment.

Method—Measures included a clinician-report of content addressed in sessions and additional measures of treatment progress and process (e.g., therapeutic alliance) completed by clinicians, clients, and their caregivers. Item responses in the top 25th percentile in severity from these measures appeared as ‘problem alerts’ on corresponding computerized feedback reports. Clinicians randomized to the feedback group received feedback weekly while the control group did not. Analyses were conducted using the Cox proportional hazards regression for recurrent events.

Results—For all content domains, the results of the survival analyses indicated a robust effect of the feedback group on addressing specific content in sessions, with feedback associated with shorter duration to first occurrence and increased likelihood of addressing or focusing on a topic compared to the non-feedback group.

Conclusion—There appears to be an important relationship between feedback and cumulative problem alerts reported by multiple informants as they influence session content.

Keywords

Child Psychotherapy; Mental Health Services Research; Technology in Psychotherapy Research & Training; Process Research; measurement feedback system; Session Report Form

Corresponding Author: Susan R. Douglas Ph.D., Department of Psychology and Human Development, Peabody College, 230 Appleton Place #552, Hobbs 211B, Nashville, TN 37203; (phone) 615-322-8141; (fax) 615-343-9494; Susan.Douglas@vanderbilt.edu.

Note: Susan Douglas has been previously published as Susan Douglas Kelley.

While not yet widespread, there is growing evidence that providing clinicians with feedback about client treatment progress improves outcomes for youth and adult clients (e.g., Bickman et al., 2011/2015; Chorpita, Bernstein, & Daleiden, 2008/2011; Knaup et al., 2009; Lambert et al., 2003/2005; Shimokawa, Lambert, & Smart, 2010). In comparison, the use of systematic feedback to improve outcomes has been a proven quality improvement tool for decades in related fields, such as management, education and health care (e.g., Kluger & DeNisi, 1996). Based on the available evidence to date, it has been recommended that greater use of outcome monitoring and feedback should occur in everyday mental health practice (APA Presidential Task Force on Evidence-Based Practice, 2006; Kazdin, 2008; Weisz, Ng, & Bearman, 2014).

Feedback is a foundational element of an evidence-based practice approach, which calls for the integration of research evidence with clinical decision-making and contextual factors associated with the client and setting (APA Task Force, 2006). The current position in psychotherapy research generally supports the benefits of routine outcome monitoring and feedback (e.g., Shimokawa et al., 2010; Youn, Kraus, & Castonguay, 2012), albeit with some caveats because the challenges in implementation may be substantial. For example, in addition to practical considerations like time burden and resources, mental health clinicians may overestimate their own abilities to assess treatment progress and outcomes based on clinical judgment leading to lower utilization of outcome management systems (Boswell, Kraus, Miller, & Lambert, 2015). Importantly, we do not yet know enough about the mechanisms of action and role-relationship implications of using routine outcome monitoring and feedback to inform performance evaluations (Evans, 2012; Noell & Gansle, 2014).

While the concept of using feedback to influence behavior makes intuitive sense, the design and implementation of feedback interventions must be approached with caution to avoid detrimental effects. For example, over a third of feedback interventions in Kluger and DeNisi's (1996) meta-analysis actually resulted in poorer performance. Comprehensive theories of feedback interventions integrate information from social and cognitive psychology, organizational theory, and management (Kluger & DeNisi, 1996; Riemer & Bickman, 2011; Riemer, Rosoff-Williams, & Bickman, 2005). However, there is little research evidence about what characteristics of feedback interventions in mental health are most effective in influencing behavior. Critical issues to explore further are feedback delivery schedules and procedures (e.g., actual mechanisms like technology or mixed technology and paper, the visual presentation like graphics, etc.), to whom feedback is provided, and the sources of data that inform the feedback (Kluger & DiNisi, 1996; Noell & Gansle, 2014). In addition, there are also client, clinician, and setting characteristics to consider, particularly as they influence implementation (de Jong, 2014).

Promising characteristics of feedback include greater frequency of feedback (weekly or bi-weekly) and the inclusion of information on treatment progress. In addition, it has been recommended that feedback is most effective when data collection is simple, flexible, and transparent and the feedback is specific, accepted as valid, and actionable (Boswell et al., 2015; Cannon & Witherspoon, 2005; Knaup et al., 2009; Riemer & Bickman, 2011). By

actionable, we mean that specific feedback elements are linked to actions that can be taken in response, as in a clinical support tool or a recommended behavior by a supervisor.

Feedback is intended to improve the accuracy of clinicians' judgments about client progress thus enabling them to better tailor treatment activities consistent with the client's progress (or lack thereof). For example, Worthen and Lambert (2007) hypothesized that feedback from clients influenced clinical outcomes by providing clinicians with information that they may have unintentionally overlooked or underemphasized while also identifying problems within domains that could jeopardize progress. This is a particularly important feature of feedback, given the poor accuracy of clinicians in predicting or identifying client deterioration (Hatfield et al., 2009; Hannan et al., 2005) and low agreement between clinicians and clients on the focus of treatment (Hawley & Weisz, 2003). As Kazdin (2008) stated, clinicians are "researchers in the sense of hypothesizing that a particular treatment combination will have particular effects and testing this hypothesis with the individual case (p. 155)." Feedback can provide systematic data to enhance, not replace, clinical decision-making.

We know little about how feedback affects clinical decision-making and clinician behavior in treatment sessions (Lambert et al., 2003) or in supervision (Sparks et al., 2011; Worthen & Lambert, 2007). Thus we can only speculate about how feedback influences outcomes. As a result of feedback clinicians may: address the feedback directly with the client; change their case conceptualization and resulting interventions to introduce into treatment; delay discharge due to new concerns; become more attentive to a client; discuss feedback within supervision sessions in order to increase awareness of the treatment process; or they may ignore the feedback altogether. What these actions have in common is that the clinician is central to linking feedback to client outcomes. Understanding how clinician behavior is influenced by feedback is an important next step in advancing the use of feedback to improve client outcomes.

For youth clients, it is unknown how clinicians may change their behavior based on feedback from different sources about the same client. While clinicians generally value the client's and family's perspectives, there can be doubts about the perceived validity of their feedback due to competing motives (e.g., an adolescent understates symptoms in order to discontinue treatment) or lack of insight into one's own problems (Riemer & Bickman, 2011; Sapyta et al., 2005). A better understanding of how session content may be correlated with problems reported by the youth and caregiver should provide some insight into how feedback influences clinician behavior.

The current study was intended as the next step to build on previous findings (Douglas Kelley et al., 2010/2012) and is part of a larger study investigating the effects of feedback to clinicians on clinician behavior and client outcomes in community youth mental health treatment. The intervention is called Contextualized Feedback Systems (CFS). The effectiveness of CFS has been established in two longitudinal randomized experiments, one where randomization occurred at the site level (Bickman et al., 2011) and other where randomization occurred at the client level (Bickman et al., 2015). Both studies found an effect of feedback on client outcomes for youths receiving home-based and outpatient

mental health services respectively. The purpose of this study was to explore how clinician-reported content addressed in treatment sessions was predicted by clinician feedback group and multi-informant cumulative problem alerts that appeared in computerized feedback reports.

The current study has three aims: to explore whether feedback impacts (1) the likelihood of addressing content in sessions per clinician's report; (2) the timeliness of addressing content in sessions; and (3) session content differentially based on the presence of dynamic problem-specific severity. Two hypotheses were proposed. First, it is anticipated that the relative likelihood of addressing content will be higher for clients whose clinicians were in the feedback condition compared to the non-feedback condition. Second, it is hypothesized that the time to addressing content will be shorter for clients whose clinicians were in the feedback condition compared to the non-feedback condition. Both of these hypotheses are predicated on the possibility of increased awareness of client problems based on viewing related problem alerts in the feedback reports. The third aim is exploratory and was partially based on previous findings from the Douglas Kelley and colleagues (2012) study, which found session content correlated differently based on type of problem and reporter.

Methods

Feedback Intervention

Contextualized Feedback Systems (CFS) has been described in detail elsewhere (Bickman et al., 2011/2014). Feedback intervention characteristics included weekly clinical feedback to clinicians on youth treatment progress (e.g., symptoms and functioning) and treatment process (e.g., therapeutic alliance) intended to alert clinicians to youth or caregiver problems they may not have perceived or ignored, to assist in assessing client progress so as to inform treatment plans and goals, and to enhance dialogue and case understanding in clinical supervision. In training and ongoing support for clinicians in the feedback group, clinicians were taught how to interpret feedback reports where problem alerts were present. Specifically, they were asked to consider the clinical utility of incorporating the feedback in the very next session by addressing any alerts that were unexpected or indicated a decline in treatment progress or process. The development and refinement of CFS was a user-centered collaboration between researchers and practitioners (Bickman, Douglas Kelley, and Athay, 2012; Riemer, Douglas Kelley, Casey, & Taylor Haynes, 2012). CFS implementation training and support was designed to emphasize integration with clinical values and workflow, promote administrative leadership, and facilitate sustainability by encouraging innovation (Douglas, Button, & Casey, 2014).

Procedures

The Peabody Treatment Progress Battery (PTPB; Bickman, Athay et al., 2010) measures were administered using paper- and pencil-forms at the close of treatment sessions. Respondents were instructed to complete measures privately, then placed into an envelope that was then sealed. Thus, the administration was designed so that problem alerts and other feedback would not be accessible except to clinicians in the feedback group when they viewed feedback reports on their computers. As soon as an administrative assistant data

entered the measures into a computer, feedback reports became automatically available for viewing by clinicians in the feedback condition. Feedback viewing was recorded within the CFS application, with approximately one-third (35%) of the clinicians in the feedback group who did not view any feedback reports. On average, clinicians in the feedback group who viewed at least one report viewed 41% (s.d. = 27%) of reports, with viewing rates ranging from 2% to 100%.

Clinicians completed the Session Report Form (SRF; Douglas Kelley et al., 2010) at the close of each session, documenting whether they addressed, focused on, or did not address SRF topics as part of that treatment encounter. In addition, all participants in the session (clients, caregivers, clinicians) completed the scheduled PTPB measures relevant to treatment progress and process resulting in item-level problem alerts corresponding to specific SRF topics. Notably, the problem alerts that were contained in feedback reports were lagged, meaning that they included all previous item alerts up to but not including that session when the SRF was completed. However, because of the possibility the completion of measures spurred conversation in the session, a conservative approach was used where the cumulative problem alert variables were created to include all previous item alerts up to and including that session. The Vanderbilt University Institutional Review Board approved all procedures with a waiver of informed consent.

Participants

Participants were drawn from the larger study that first evaluated the effects of CFS on youth outcomes (Bickman et al., 2011). They are the same sub-sample used in the previous study (Douglas Kelley et al., 2012) that explored the relationship between clinician-reported session content and concurrent multi-informant problem alerts. All clients had presenting problems encompassing a range of common youth mental health issues (e.g., mood and anxiety disorders, school problems, oppositional behavior, and impulsivity/attention deficit disorders) and were receiving 'treatment as usual' mental health care, typically in home-based settings.

Twenty-eight sites in ten states were included in the larger study, with randomization to the feedback or non-feedback conditions occurring at the site level. There were no baseline differences between the feedback and non-feedback conditions on demographic and background characteristics for clinicians, with one significant finding for youths and caregivers (there were more black and fewer white youths and caregivers in the feedback group compared to the control condition; Bickman et al., 2011). To be included in the current study, youths from the larger evaluation sample had to be new clients receiving home-based counseling with at least one session where the clinician completed the SRF measure. A total of 299 clients aged 11 to 18 years were included, with a total of 3,143 sessions with SRFs completed by 133 clinicians. Table 1 shows client's background characteristics by randomized condition. Similar to the initial evaluation study there were no significant differences at the adjusted p-value ($p < 0.01$) with the exception of client and caregiver race (see Bickman et al., 2011 for more detail on caregiver and clinician background characteristics).

Measures

The Session Report Form (SRF; Douglas Kelley et al., 2010) is a 25-item self-report measure developed collaboratively by a team of clinicians and researchers. Previous analyses of the SRF indicate some evidence of a distinct subscale of treatment process using principal components analysis (Cronbach's $\alpha = 0.81$). There is also evidence of content validity with an adequate range of topics related to session content as well as discrimination in patterns of client and clinician topics addressed. The SRF was completed by clinicians as part of the documentation associated with a specific treatment session (e.g., progress notes), and is intended to be completed at the end of each session. It is part of the larger *Peabody Treatment Progress Battery* (PTPB) of measures of youth treatment process and progress (Bickman, Athay, et al., 2010). In addition to items describing session characteristics (i.e., who participated, length and location of the session, rating of the session, and time spent in crisis) the SRF contains 20 content domains that could be addressed as part of the session. For each content domain (e.g., emotional issues, family issues) the clinician reported whether the topic was not addressed, addressed in, or a focus of the session. For this study, responses were recoded as binary variables (not addressed v addressed/focus).

Almost all topics corresponded to other measures in the PTPB. For example, the SRF topic "emotional issues" is a content domain that related to 13 items related to internalizing symptoms reported by the youth, caregiver and clinician on the measure of youth symptoms and functioning shown in Table 2. All of the PTPB measures have excellent psychometric properties with strong evidence of reliability, evidence for construct validity including estimates of factorial and discriminant validity for all measures and concurrent validity for the Symptoms and Functioning Severity Scale (SFSS) (Bickman et al., 2010; Bickman & Athay, 2012). Internal consistency statistics are provided in the footnotes for Table 2. Table 2 presents the PTPB measures that were sources for multi-informant problem alerts. Items were rated as 'problem alerts' when the response was in the top 25% of the distribution of item severity from the psychometric sample (Bickman et al., 2010). Problem alerts are item and informant specific. Thus, where there were multiple informants of a measure, whether or not an item counted as a problem alert could differ by respondent.

Analytic Approach

Data were available at the session level. Because feedback reports included item alerts prior to the session, the analytic models had to account for cumulative item alerts as a lagged time-varying covariate. Cumulative item alerts was coded as 1 where one or more item alerts existed for sessions prior and up to that particular SRF administration and 0 for no previous item alerts. The other independent variable, feedback group was constant, coded as 0 for the non-feedback condition and 1 for the feedback condition. The interaction term was the product of multiplying the two variables. The dependent variables of interest were the 18 SRF topic variables that were coded as 0 for not addressed and 1 for addressed in or a focus of the session. Although clinicians could discriminate topics that were simply addressed from those that were a focus of the session, these responses were combined to create a binary variable to simplify the statistical models.

Since the dependent variable could be considered as recurrent events that may or may not occur repeatedly throughout the observation period, we used survival analysis of repeated events (Hosmer, Lemeshow, & May, 2008; Twisk, Smidt, & de Vente, 2005). The estimation method used was the Cox proportional hazards for recurrent events model, which is a common method used to analyze time to success or failure (i.e., occurrence of the event) separately for repeated events, adjusted for the clustering of time periods within clients. This model is a multivariate approach and treats each recurrent event as identical. The counting approach was used, where each time period was analyzed separately with the assumption that clients were at risk each time irrespective of whether the event occurred in the prior time period (Twisk et al., 2005). The robust standard error was reported to account for the nesting of recurrent events within clients. Each model results in output that includes a regression coefficient and a hazard ratio. The regression coefficient represents the duration to occurrence of a failure event, defined here as the topic being addressed/focused upon in the session¹. A positive coefficient means a shorter duration and a negative coefficient means a longer duration. The hazard ratio is the exponent of the coefficient, with hazard ratios above 1.0 representing increased likelihood of addressing/focusing on a topic and those below 1.0 representing reduced likelihood.

For each independent variable, five models were estimated. The first was the simple model, which estimated the relative likelihood of addressing/focusing on a topic based on randomized condition of clinicians to the feedback group or the non-feedback group. The second was the adjusted model with two covariates, which estimated the relative likelihood of addressing/focusing on a topic for feedback group adjusting for cumulative problem alerts and vice versa. The third was the interaction model with the two covariates and their interaction. The null group is the non-feedback condition with a value of 0 (feedback group = 0 times cumulative problem alerts = 0). The treatment group is the feedback condition with a value equal to the cumulative problem alerts (feedback group = 1 times cumulative problem alerts = cumulative problem alerts). As per Hosmer, Lemeshow, and May's (2008) recommendations, the three models were compared and the most complete model that reached significance was chosen for interpretation (shaded in grey in the tables in the results section below). The 2nd and 3rd models were run with two sets of variables (clinician-reported problem alerts, combined client/caregiver-reported problem alerts) for a total of five models per variable. To adjust for the number of models, the p value was adjusted to $p < .05/5$ or $p < 0.01$. Kaplan-Meier plots were produced to aid in interpreting the regression coefficients that were used to determine the effect of feedback on time to first addressing/focusing on a topic.

Results

Table 3 presents descriptive information for all clients on both session content that was addressed or focused upon in sessions and cumulative problem alerts (reported by the clinician and the combined client/caregiver) by SRF topic. As can be seen in the first

¹In survival analyses most commonly reported in the medical literature, the failure event is a negative one, such as death, relapse, or other treatment failure. In this study, the failure event is positive (the addressing or focusing on a topic). Thus, the desired direction of effect is positive (we want the event to happen) rather than what is typically seen in the medical literature (where the failure event is not desired).

column on session content by topic, there were very few clients with missing data points for the dependent variable (1 to 3% of clients had a missing SRF topic). The frequency of session content being addressed/focused on in sessions varied by topic. Of note, there was wide variation in missing data for problem alerts. While the large majority of clients (90-94% depending on reporter) had data for the first 9 topics that corresponded to the SFSS, the percentage of clients with non-missing data for treatment process and caregiver-related topics (the last 9 topics) ranged from 32% to 90%. This was largely because those measures were typically scheduled less frequently (see Table 2).

A series of independent samples *t*-tests (equal variances not assumed) were conducted to determine whether there were any significant differences by feedback group for the cumulative problem-alert variables. None approached significance at the adjusted *p*-value < 0.01.

The regression coefficients were significantly positive ($p < 0.05$) for almost all topics with the exception of problems with delinquent behavior and alcohol/ substance use. This means that for 16 of the 18 topics, the time to addressing content was shorter for clients whose clinicians were in the feedback condition compared to the non-feedback condition. To provide a visual example, Figure 1 shows the Kaplan-Meier survival estimates for the topic of behavioral issues. The graph shows that clients whose clinicians were in the feedback group have a lower survival rate than clients whose clinicians were in the control condition. In other words, the duration to event is shorter. For example, at 20 weeks the survival rate is 0.25 for the feedback group and 0.35 for the non-feedback group, which means behavioral issues have been addressed/focused on for 75% of the clients in the feedback group compared to 65% of the clients in the non-feedback group. For the two exceptions (problems with delinquent behavior and alcohol/ substance use) where feedback did not shorten time to addressing content, both were typically addressed infrequently (for about half the clients) with the first occurrence after several weeks in treatment (average time was about seven weeks).

There was a significant effect of feedback on the likelihood of addressing and/or focusing on every SRF topic. A summary of the significant model results from the survival analyses can be found in Figure 2 (note that the results for all models are presented in the appendix). As recommended (Hosmer, Lemeshow, & May, 2008), we compared the three models per reporter per topic and chose the most complete model for interpretation that was significant at an adjusted *p*-value of $p < 0.01$. The figure shows the hazard ratio and the 95% confidence interval for each selected model by reporter (clinician, client and/or caregiver combined) and SRF topic. All hazard ratios were above 1.0, indicating significantly increased likelihood of addressing/focusing on a topic. Hazard ratios can be interpreted by subtracting 1.0 then converting the result into the percent increased likelihood. In order of increasing complexity, the three models were the simple (S) model, the adjusted (A) model, and the interaction (I) model.

The simple model is most consistent with the first hypothesis that the relative likelihood of addressing content will be higher for clients whose clinicians were in the feedback condition compared to the non-feedback condition. Six of the SRF topics were best explained by this

model. For example, for client-reported perceptions of counseling impact the simple model was the best explanatory model, indicating that clinicians in the feedback group were on average 66% (Hazard ratio of 1.66) more likely to address/focus on client-reported counseling impact problems than clinicians in the control condition. Other topics where the simple model provided the best explanation were clinician-rated family issues, client and/or caregiver-rated emotional issues, client-rated satisfaction with services, caregiver strain, and caregiver satisfaction with services (respectively 33%, 39%, 65%, 95%, and 134% more likely to be addressed/focused on by clinicians in the feedback group than the control condition).

There were six topics where the adjusted model provided the best explanatory fit, indicating that feedback increased the relative likelihood of addressing or focusing on a topic even after adjusting for cumulative problem alerts. For example, after adjusting for clinician-reported cumulative problem alerts related to client harm to self or others, the model in Figure 2 indicates that clinicians in the feedback group were 64% more likely to address/focus on harm as a topic in the next session than were clinicians in the non-feedback condition. The other topics all included combined client/caregiver-reported problem alerts in the adjusted models as shown in Figure 2, including family issues, client hope for the future, therapeutic relationship with the client and the caregiver separately, and caregiver satisfaction with life (respectively 46%, 24% 63% 73%, and 104% more likely).

The interaction between feedback group and problem alerts provided the best explanatory model for eight of the topics (a total of 13 of the models differing by reporter). For all of these topics, the interaction of cumulative problem alerts heightened the impact of feedback. To interpret the interaction models, the hazard ratio presented in Figure 2 represent the sum of hazard ratios for feedback plus the interaction term where there were cumulative problem alerts present prior to the clinician's completion of the SRF. For example, the alcohol/substance use topic is significant for the interaction model for both clinician-reported and combined client/caregiver-reported problem alerts (hazard ratio for the feedback group were 2.22 and 2.15 respectively). This indicates that for clients with reported alcohol or substance use problems, the clinicians in the feedback condition were over two times more likely to address and/or focus on this issue in a following session than were clinicians in the control condition.

Discussion

This study represents an important next step in exploring how feedback provided to clinicians may influence their behavior in sessions with youths receiving treatment in community mental health settings. The findings show that the duration to first occurrence and relative likelihood of clinicians addressing content related to specific domains of youth treatment progress and process was significantly improved for clinicians randomized to the feedback group compared to the non-feedback group. For a few topics, there appears to be a simple relationship between feedback group and increased likelihood of addressing or focusing on content. However, for most content areas there is an important relationship between feedback group and the cumulative problem alerts reported by multiple informants. In other words, the effect of feedback group is enhanced by the presence of problem alerts.

Previous research on the SRF found that clinicians had a general style in their approach to treatment, accounting for more of the variance than did clients on several topics related to youth symptoms and functioning (Douglas Kelley et al., 2010). A follow-up study on the relationship between clinician-reported content addressed in sessions and concurrent multi-informant problem alerts found several differences in session content based on type of problem alert and reporter. Clinicians reported addressing content related to emotional, family, and friend/peer problems more when reported by youths, while content related to youth behavioral problems occurred more when reported by caregivers (Douglas Kelley et al., 2012). High-risk problems (alcohol/substance use, harm to self or others) were strongly correlated with session content when reported by anyone (clinician, youth, or caregiver). Problems with the treatment process (e.g., therapeutic alliance) and caregiver strain were not associated with session content regardless of reporter.

The results presented here showed that several treatment process- and caregiver-related content domains do increase in likelihood in relation to feedback group, but are not influenced by matching problem alerts (e.g., the simple model). It is possible that there is an activating effect of feedback, where clinicians become more generally attuned to issues of importance to clients and their caregivers. Further, it is clear that the client and the caregiver are bringing issues of importance to the fore through their completion of PTPB measures. Notably, there were no differences in the distribution of problem alerts by treatment condition between clients whose clinicians were in the feedback group compared to the non-feedback group. Yet, there remained an effect of feedback on the likelihood of addressing and/or focusing on those problems either after accounting for cumulative problem alerts or including it as an interaction term in the models.

Although this study did not include outcome data in analyses, there are two separate randomized clinical trials of the same feedback system (one with the same dataset) that both show faster improvements in youth symptoms and functioning for clients whose clinicians were in the feedback group compared to the control condition (Bickman et al., 2011/2015). The present study showed that for all but two of the content areas, clinicians in the feedback group addressed topics earlier in treatment than did clinicians in the control group. In regard to clinical significance, hazard ratios are used to express the effects associated with the relative differences between treatment conditions (Hosmer, Lemeshow, & May, 2008), which can be more useful than simply considering statistical significance. Nine of the SRF topics had hazard ratios between 2.15 and 5.11, which can be interpreted as clinicians in the feedback group being 115% to 411% more likely to address a topic than clinicians in the comparison group. Hazard ratios for the remaining nine topics ranged between 1.33 and 1.95, indicating smaller but still potentially clinically meaningful effects.

The practice implications suggested by the study findings seem clear. Routine outcome monitoring and feedback may increase clinicians' responsiveness to common problems reported by clients and caregivers in youth psychotherapy, at least in terms of self-reported content addressed in sessions. This is consistent with a common factors approach to psychotherapy, where non-model specific factors such as motivation for treatment or the therapeutic alliance have been found to be associated with improved outcomes (e.g., Lambert & Barley, 2002; Wampold, 2001). It is also consistent with the research on client-

centered clinician communication behaviors, which typically include active use of the client's own language for problems and treatment goals and partnering in collaborative treatment planning (Levinson, Lesser, & Epstein, 2010; Makoul, 2001).

This is not an argument against specific treatment modalities (e.g., cognitive-behavioral therapy). Common factors of treatment and client-centered communication interventions can be combined with specific treatment modalities (e.g., cognitive-behavioral therapy) to enhance outcomes. Rather, these findings provide evidence to support the importance of a broad-based common factors approach to the selection of measures for inclusion in routine outcome monitoring and feedback (Douglas Kelley, Norwood, & Bickman, 2010). The present study was conducted in community mental health centers with 'treatment as usual'. It would be interesting to explore how the effects of feedback on addressing session content changed in the context of an evidence-based treatment, where fidelity to the treatment model is valued.

In an effort to gain greater understanding of clinician behavior in sessions, several researchers have focused on specific treatment strategies or techniques (Bearsley-Smith et al., 2008; Chorpita et al., 2008; Chorpita, Bernstein, & Daleiden, 2011; Garland et al., 2009). An alternate approach was taken here was to focus on content rather than techniques, since content is more universally understood and less likely to be subject to misinterpretation or require extensive training for form completion. As a clinician self-reported measure, the Session Report Form (SRF) holds practical value in ease of use; however, other methods of measuring content of sessions may be more accurate, such as observational coding systems. Hurlbert and colleagues (2010) compared ratings of the occurrence and intensity of clinical strategies as reported by clinicians to standardized coding of session videotapes. They found that clinicians were 2.5 times more likely to overestimate the number of goals and strategies pursued in each session compared to trained observers. Further, correspondence on specific goals and strategies was low. It is quite possible that clinicians may self-report intended behaviors that may not be perceptible to others.

It should be noted that these session-based data are highly complex with multiple measures, multiple informants, and time-varying effects on clinician-reported content addressed in sessions. While the recurrent events survival analysis techniques used here offer a sophisticated approach to dealing with this type of data structure, two analytic limitations remain to be explored. First, while the use of the robust standard error accounted for clustering of events within clients (Twisk et al., 2005) these models did not adjust for clustering of clients within clinicians. Second, the Cox proportional hazards regression model assumes that recurrent events are identical, resulting in the hazard ratio being constant. The stratified Cox model is an alternative model that can address this issue (Hosmer, Lemeshow, & May, 2008).

This study was intended to investigate a vast, complex dataset that has been collected in a real-life clinical setting. While the results provide evidence supporting the hypothesis that one mechanism of action for feedback to clinicians about clients' problems is that it increases the likelihood of addressing these problems in sessions, further research is needed.

In addition to the analytic limitations mentioned here, there are also important relationships that were not modeled. There are potential client, setting, and clinician characteristics that may influence the impact of feedback on addressing or focusing on session content. In addition, we know that a third of the clinicians in the feedback group did not view a single feedback report for any of their clients (Bickman et al., 2011). Both of the evaluation studies linked to this study have shown a strong dose-response relationship between implementation (defined as rates of measure completion and feedback viewing). This would be an important area to explore since better implementation may improve even further the relationship between feedback, multi-informant problem alerts, and addressing/focusing on session content.

Acknowledgments

Preparation of this article was partially supported by the National Institute of Mental Health of the National Institutes of Health under Award Number R01087814. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Appendix

Table 1
Results of Cox regression for recurrent events for simple, adjusted, and interaction models where addressing content was predicted by feedback group, clinician reported dynamic problem-specific severity, and their interaction

Topic	Survival Analysis Model	Reg. coeff. (b)	S.E. (b)	p-value	Hazard ratio	95% CI for hazard ratio
Emotional issues	<i>Simple model</i> - Feedback group	0.33	0.04	0.000	1.39	1.29, 1.51
	<i>Adjusted model</i> - Feedback group			0.000	1.44	1.33, 1.57
	Cumulative problem alerts (clinician)			0.086	1.07	0.99, 1.15
	<i>Interaction model</i> - Feedback group			0.086	1.93	0.91, 4.09
	Cumulative problem alerts (clinician)			0.231	1.05	0.97, 1.15
	Feedback * Cumulative alerts			0.008	1.01	1.00, 1.02
Behavioral issues	<i>Simple model</i> - Feedback group	0.44	0.04	0.000	1.55	1.43, 1.68
	<i>Adjusted model</i> - Feedback group			0.000	1.61	1.48, 1.75
	Cumulative problem alerts (clinician)			0.000	1.05	1.03, 1.08
	<i>Interaction model</i> - Feedback group			0.000	4.10	2.50, 6.74
	Cumulative problem alerts (clinician)			0.009	1.04	1.01, 1.06

Topic	Survival Analysis Model	Reg. coeff. (b)	S.E. (b)	p-value	Hazard ratio	95% CI for hazard ratio
	Feedback * Cumulative alerts			0.001	1.01	1.01, 1.02
Family issues	<i>Simple model</i> - Feedback group	0.29	0.04	0.000	1.33	1.23, 1.44
	<i>Adjusted model</i> - Feedback group			0.000	1.36	1.26, 1.48
	Cumulative problem alerts (clinician)			0.827	1.00	0.99, 1.01
	<i>Interaction model</i> - Feedback group			0.948	1.00	0.83, 1.20
	Cumulative problem alerts (clinician)			0.841	1.00	0.99, 1.01
	Feedback * Cumulative alerts			0.183	1.01	1.00, 1.02
School/work issues	<i>Simple model</i> - Feedback group	0.22	0.04	0.000	1.25	1.15, 1.36
	<i>Adjusted model</i> - Feedback group			0.000	1.28	1.17, 1.39
	Cumulative problem alerts (clinician)			0.007	1.03	1.01, 1.05
	<i>Interaction model</i> - Feedback group			0.604	0.93	0.71, 1.22
	Cumulative problem alerts (clinician)			0.074	1.02	1.00, 1.04
	Feedback * Cumulative alerts			0.000	1.02	1.01, 1.03
Friends/peer issues	<i>Simple model</i> - Feedback group	0.15	0.04	0.000	1.17	1.07, 1.27
	<i>Adjusted model</i> - Feedback group			0.000	1.23	1.12, 1.34
	Cumulative problem alerts (clinician)			0.376	1.02	0.98, 1.05
	<i>Interaction model</i> - Feedback group			0.168	1.49	0.85, 2.61
	Cumulative problem alerts (clinician)			1.000	1.00	0.96, 1.04
	Feedback * Cumulative alerts			0.000	1.02	1.01, 1.03
Problems w/delinquent behavior	<i>Simple model</i> - Feedback group	0.18	0.08	0.027	1.20	1.02, 1.41
	<i>Adjusted model</i> - Feedback group			0.060	1.17	0.99, 1.39
	Cumulative problem alerts (clinician)			0.011	1.04	1.01, 1.07
	<i>Interaction model</i> - Feedback group			0.277	0.78	0.50, 1.22
	Cumulative problem alerts (clinician)			0.208	1.02	0.99, 1.05

Topic	Survival Analysis Model	Reg. coeff. (b)	S.E. (b)	p-value	Hazard ratio	95% CI for hazard ratio
	Feedback * Cumulative alerts			0.000	1.06	1.04, 1.07
Alcohol/ substance use	<i>Simple model</i> - Feedback group	-0.01	0.09	0.904	0.99	0.83, 1.17
	<i>Adjusted model</i> - Feedback group			0.850	0.99	0.82, 1.17
	Cumulative problem alerts (clinician)			0.508	1.01	0.99, 1.03
	<i>Interaction model</i> - Feedback group			0.313	1.17	0.87, 1.58
	Cumulative problem alerts (clinician)			0.213	0.99	0.97, 1.01
	Feedback * Cumulative alerts			0.000	1.05	1.03, 1.07
Harm to self or others	<i>Simple model</i> - Feedback group	0.43	0.09	0.000	1.54	1.30, 1.82
	<i>Adjusted model</i> - Feedback group			0.000	1.64	1.37, 1.96
	Cumulative problem alerts (clinician)			0.000	0.96	0.95, 0.98
	<i>Interaction model</i> - Feedback group			0.000	1.71	1.28, 2.28
	Cumulative problem alerts (clinician)			0.000	0.96	0.94, 0.98
	Feedback * Cumulative alerts			0.855	1.00	0.98, 1.02

Note: significance determined at adjusted p-value of $p < 0.01$

Table 2
Results of Cox regression for recurrent events for simple, adjusted, and interaction models where addressing content was predicted by feedback group, client and/or caregiver reported dynamic problem-specific severity, and their interaction

Variable	Topic	Reg. coeff. (b)	S.E. (b)	p-value	Hazard ratio	95% CI for hazard ratio
Emotional issues	<i>Simple model</i> - Feedback group	0.33	0.04	0.000	1.39	1.29, 1.51
	<i>Adjusted model</i> - Feedback group			0.000	1.48	1.36, 1.61
	Cumulative problem			0.809	1.10	0.93, 1.10

Variable	Topic	Reg. coeff. (b)	S.E. (b)	p-value	Hazard ratio	95% CI for hazard ratio
	alerts (cl / cg)					
	<i>Interaction model</i> - Feedback group			0.000	4.72	2.08, 10.7
	Cumulative problem alerts (cl / cg)			0.280	0.95	0.87, 1.04
	Feedback * Cumulative alerts			0.025	1.01	1.00, 1.02
Behavioral issues	<i>Simple model</i> - Feedback group	0.44	0.04	0.000	1.55	1.43, 1.68
	<i>Adjusted model</i> - Feedback group			0.000	1.61	1.48, 1.75
	Cumulative problem alerts (cl / cg)			0.032	1.06	1.01, 1.11
	<i>Interaction model</i> - Feedback group			0.559	0.83	0.45, 1.54
	Cumulative problem alerts (cl / cg)			0.259	1.03	0.98, 1.09
	Feedback * Cumulative alerts			0.005	1.01	1.00, 1.02
Family issues	<i>Simple model</i> - Feedback group	0.29	0.04	0.000	1.33	1.23, 1.44
	<i>Adjusted model</i> - Feedback group			0.000	1.46	1.34, 1.59
	Cumulative problem alerts (cl / cg)			0.000	1.01	1.00, 1.02
	<i>Interaction model</i> - Feedback group			0.000	2.00	1.53, 2.62

Variable	Topic	Reg. coeff. (b)	S.E. (b)	p-value	Hazard ratio	95% CI for hazard ratio
	Cumulative problem alerts (cl / cg)			0.789	1.00	0.99, 1.01
	Feedback * Cumulative alerts			0.124	1.01	1.00, 1.02
School/work issues	<i>Simple model</i> - Feedback group	0.22	0.04	0.000	1.25	1.15, 1.36
	<i>Adjusted model</i> - Feedback group			0.000	1.41	1.29, 1.54
	Cumulative problem alerts (cl / cg)			0.001	1.02	1.01, 1.04
	<i>Interaction model</i> - Feedback group			0.001	1.80	1.27, 2.54
	Cumulative problem alerts (cl / cg)			0.063	1.01	1.00, 1.03
	Feedback * Cumulative alerts			0.000	1.02	1.01, 1.02
Friends/peer issues	<i>Simple model</i> - Feedback group	0.15	0.04	0.000	1.17	1.07, 1.27
	<i>Adjusted model</i> - Feedback group			0.000	1.27	1.16, 1.39
	Cumulative problem alerts (cl / cg)			0.837	1.00	0.98, 1.03
	<i>Interaction model</i> - Feedback group			0.242	1.40	0.80, 2.44
	Cumulative problem alerts (cl / cg)			0.669	1.00	0.97, 1.02
	Feedback *			0.000	1.02	1.01, 1.03

Variable	Topic	Reg. coeff. (b)	S.E. (b)	p-value	Hazard ratio	95% CI for hazard ratio
	Cumulative alerts					
Problems w/delinquent behavior	<i>Simple model</i> - Feedback group	0.18	0.08	0.027	1.20	1.02, 1.41
	<i>Adjusted model</i> - Feedback group			0.002	1.31	1.10, 1.55
	Cumulative problem alerts (cl / cg)			0.000	1.12	1.05, 1.19
	<i>Interaction model</i> - Feedback group			0.269	0.74	0.43, 1.27
	Cumulative problem alerts (cl / cg)			0.005	1.09	1.03, 1.16
	Feedback * Cumulative alerts			0.000	1.05	1.03, 1.06
Alcohol/ substance use	<i>Simple model</i> - Feedback group	-0.01	0.09	0.904	0.99	0.83, 1.17
	<i>Adjusted model</i> - Feedback group			0.159	1.14	0.95, 1.37
	Cumulative problem alerts (cl / cg)			0.016	1.03	1.01, 1.06
	<i>Interaction model</i> - Feedback group			0.619	1.09	0.79, 1.50
	Cumulative problem alerts (cl / cg)			0.363	1.01	0.99, 1.04
	Feedback * Cumulative alerts			0.000	1.06	1.04, 1.07
Client hope for future	<i>Simple model</i> - Feedback group	0.16	0.05	0.001	1.18	1.07, 1.29

Variable	Topic	Reg. coeff. (b)	S.E. (b)	p-value	Hazard ratio	95% CI for hazard ratio
	<i>Adjusted model</i> - Feedback group			0.001	1.24	1.10, 1.41
	Cumulative problem alerts (client)			0.000	0.98	0.97, 0.99
	<i>Interaction model</i> - Feedback group			0.000	1.39	1.19, 1.62
	Cumulative problem alerts (client)			0.000	0.98	0.96, 0.99
	Feedback * Cumulative alerts			0.230	0.98	0.97, 1.01
Client motivation for treatment	<i>Simple model</i> - Feedback group	0.32	0.06	0.000	1.38	1.24, 1.54
	<i>Adjusted model</i> - Feedback group			0.000	1.53	1.30, 1.79
	Cumulative problem alerts (client)			0.000	1.04	1.02, 1.05
	<i>Interaction model</i> - Feedback group			0.016	1.27	1.05, 1.54
	Cumulative problem alerts (client)			0.000	1.03	1.02, 1.05
	Feedback * Cumulative alerts			0.000	1.06	1.03, 1.10
Therapeutic relationship w/client	<i>Simple model</i> - Feedback group	0.44	0.07	0.000	1.55	1.34, 1.79
	<i>Adjusted model</i> - Feedback group			0.000	1.63	1.41, 1.90
	Cumulative problem			0.006	1.02	1.01, 1.04

Variable	Topic	Reg. coeff. (b)	S.E. (b)	p-value	Hazard ratio	95% CI for hazard ratio
	alerts (client)					
	<i>Interaction model</i> - Feedback group			0.000	2.89	2.25, 3.70
	Cumulative problem alerts (client)			0.089	1.02	1.00, 1.03
	Feedback * Cumulative alerts			0.242	0.99	0.97, 1.01
Client perceptions on counseling impact	<i>Simple model</i> - Feedback group	0.50	0.06	0.000	1.66	1.47, 1.87
	<i>Adjusted model</i> - Feedback group			0.000	2.14	1.82, 2.52
	Cumulative problem alerts (client)			0.171	0.99	0.97, 1.01
	<i>Interaction model</i> - Feedback group			0.000	2.79	2.15, 3.62
	Cumulative problem alerts (client)			0.333	0.99	0.98, 1.01
	Feedback * Cumulative alerts			0.065	0.98	0.96, 1.00
Client satisfaction w/services	<i>Simple model</i> - Feedback group	0.50	0.07	0.000	1.65	1.45, 1.87
	<i>Adjusted model</i> - Feedback group			0.000	1.91	1.61, 2.26
	Cumulative problem alerts (client)			0.030	0.98	0.96, 1.00
	<i>Interaction model</i> - Feedback group			0.000	1.89	1.57, 2.28

Variable	Topic	Reg. coeff. (b)	S.E. (b)	p-value	Hazard ratio	95% CI for hazard ratio
	Cumulative problem alerts (client)			0.587	1.00	0.97, 1.02
	Feedback * Cumulative alerts			0.054	0.96	0.92, 1.00
Caregiver strain	<i>Simple model</i> - Feedback group	0.67	0.06	0.000	1.95	1.72, 2.20
	<i>Adjusted model</i> - Feedback group			0.000	2.05	1.71, 2.46
	Cumulative problem alerts (caregiver)			0.050	1.03	1.00, 1.06
	<i>Interaction model</i> - Feedback group			0.126	1.49	0.89, 2.47
	Cumulative problem alerts (caregiver)			0.058	1.03	1.00, 1.06
	Feedback * Cumulative alerts			0.484	1.00	0.99, 1.02
Caregiver satisfaction w/life	<i>Simple model</i> - Feedback group	0.78	0.08	0.000	2.18	1.88, 2.53
	<i>Adjusted model</i> - Feedback group			0.000	2.04	1.66, 2.52
	Cumulative problem alerts (caregiver)			0.000	1.02	1.00, 1.04
	<i>Interaction model</i> - Feedback group			0.000	2.57	1.83, 3.61
	Cumulative problem alerts (caregiver)			0.513	1.01	0.98, 1.03
	Feedback *			0.125	1.02	1.00, 1.04

Variable	Topic	Reg. coeff. (b)	S.E. (b)	p-value	Hazard ratio	95% CI for hazard ratio
	Cumulative alerts					
Caregiver motivation for youth treatment	<i>Simple model</i> - Feedback group	0.63	0.07	0.000	1.88	1.65, 2.15
	<i>Adjusted model</i> - Feedback group			0.000	2.17	1.76, 2.69
	Cumulative problem alerts (caregiver)			0.017	1.02	1.00, 1.04
	<i>Interaction model</i> - Feedback group			0.000	2.52	1.93, 3.29
	Cumulative problem alerts (caregiver)			0.953	1.00	0.98, 1.02
	Feedback * Cumulative alerts			0.000	1.07	1.03, 1.11
Therapeutic relationship w/caregiver	<i>Simple model</i> - Feedback group	0.73	0.10	0.000	2.08	1.71, 2.53
	<i>Adjusted model</i> - Feedback group			0.000	1.73	1.38, 2.16
	Cumulative problem alerts (caregiver)			0.004	1.04	1.01, 1.07
	<i>Interaction model</i> - Feedback group			0.111	0.78	0.58, 1.06
	Cumulative problem alerts (caregiver)			0.017	1.04	1.01, 1.07
	Feedback * Cumulative alerts			0.034	1.03	1.00, 1.05
Caregiver satisfaction w/services	<i>Simple model</i> - Feedback group	0.85	0.08	0.000	2.34	1.99, 2.75

Variable	Topic	Reg. coeff. (b)	S.E. (b)	p-value	Hazard ratio	95% CI for hazard ratio
	<i>Adjusted model - Feedback group</i>			0.000	2.83	2.18, 3.66
	Cumulative problem alerts (caregiver)			0.576	1.01	0.97, 1.06
	<i>Interaction model - Feedback group</i>			0.000	3.00	2.31, 3.89
	Cumulative problem alerts (caregiver)			0.410	1.07	0.91, 1.26
	Feedback * Cumulative alerts			0.775	0.98	0.82, 1.16

Note: significance determined at adjusted p-value of $p < 0.01$

References

- APA Presidential Task Force on Evidence-Based Practice. Evidence-based practice in psychology. *American Psychologist*. 2006; 61:271–285. [PubMed: 16719673]
- Bickman, L.; Athay, MM., editors. Administration and Policy in Mental Health and Mental Health Services Research. 2012. Youth Mental Health Measurement. special issue
- Bickman, L.; Athay, MM.; Riemer, M.; Lambert, EW.; Kelley, SD.; Breda, C.; Tempesti, T.; Dew-Reeves, SE.; Brannan, AM.; Vides de Andrade, AR., editors. Manual of the Peabody Treatment Progress Battery. 2nd. Nashville, TN: Vanderbilt University; 2010.
- Bickman L, Douglas Kelley S, Athay M. The technology of measurement feedback systems. *Journal of Couple & Family Psychology: Research and Practice*. 2012; 1(4):274–284. [PubMed: 24066274]
- Bickman L, Douglas Kelley S, Breda C, de Andrade AR, Riemer M. Effects of routine feedback to clinicians on mental health outcomes of youths: Results of a randomized trial. *Psychiatric Services*. 2011; 62:1423–1429. [PubMed: 22193788]
- Bickman L, Douglas S, Vides de Andrade AR, Tomlinson M, Gleacher A, Olin S. Implementing a measurement feedback system: A tale of two sites. Manuscript accepted for publication. 2015
- Bearsley-Smith C, Sellick K, Chesters J, Francis K. Treatment content in child and adolescent mental health services: Development of the treatment recording sheet. *Administration and Policy in Mental Health and Mental Health Services Research*. 2008; 35(5):423–435. [PubMed: 18679789]
- Boswell JF, Kraus DR, Miller SD, Lambert MJ. Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy research*. 2015; 25(1):6–19. [PubMed: 23885809]
- Cannon MD, Witherspoon R. Actionable feedback: Unlocking the power of learning and performance improvement. *The Academy of Management Executive*. 2005; 19(2):120–134.
- Chorpita BF, Bernstein AD, Daleiden EL. the Research Network on Youth Mental Health. Driving with roadmaps and dashboards: Using information resources to structure the decision models in service organizations. *Administration and Policy in Mental Health and Mental Health Services Research*. 2008; 35:114–123. [PubMed: 17987376]

- Chorpita BF, Bernstein AD, Daleiden EL. Empirically guided coordination of multiple evidence-based treatments: An illustration of relevance mapping in children's mental health services. *Journal of Consulting and Clinical Psychology*. 2011; 79:470–480. [PubMed: 21787050]
- de Jong K. Deriving effective implementation strategies for outcome monitoring from feedback theory. *Administration and Policy in Mental Health*. 2014
- Douglas S, Button S, Casey SE. Implementing for sustainability: Promoting use of a measurement feedback system for innovation and quality improvement. *Administration and Policy in Mental Health and Mental Health Services Research*. 2014
- Douglas Kelley S, Bickman L, Norwood E. Evidence-based treatments and common factors in youth psychotherapy. In: Duncan, BL.; Miller, SD.; Wampold, BE.; Hubble, MA., editors. *The heart and soul of change: What works in therapy*. 2nd. Washington, DC: American Psychological Association; 2010. p. 325–355.
- Douglas Kelley S, Vides de Andrade AR, Bickman L, Robin AV. The Session Report Form (SRF): Are clinicians addressing concerns reported by youth and caregivers? *Administration and Policy in Mental Health and Mental Health Services Research*. 2012; 39(1-2):133–145. [PubMed: 22407564]
- Douglas Kelley S, Vides de Andrade AR, Sheffer E, Bickman L. Exploring the black box: Measuring youth treatment process and progress in usual care. *Administration and Policy in Mental Health and Mental Health Services Research*. Special Issue: Psychotherapy Practice in Usual Care. 2010; 37(3):287–300.
- Evans C. Cautionary notes on power steering for psychotherapy. *Canadian Psychology/Psychologie Canadienne*. 2012; 53(2):131–139.
- Garland AF, Hurlburt MS, Brookman-Frazee L, Taylor RM, Accurso EC. Methodological challenges of characterizing usual care psychotherapeutic practice. *Administration and Policy in Mental Health and Mental Health Services*. 2009; 37
- Hannan C, Lambert MJ, Harmon C, Nielson S, Smart D, Shimokawa K, et al. A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology*. 2005; 61(2):155–163. [PubMed: 15609357]
- Hatfield D, McCullough L, Frantz SHB, Krieger K. Do we know when our clients get worse? An investigation of therapists' ability to detect negative client change. *Clinical Psychology and Psychotherapy*. 2009; 17(1):25–32. [PubMed: 19916162]
- Hawley KM, Weisz JR. Child, parent and therapist (dis) agreement on target problems in outpatient therapy: The therapist's dilemma and its implications. *Journal of Consulting and Clinical Psychology*. 2003; 71(1):62–70. [PubMed: 12602426]
- Hosmer, DW.; Lemeshow, S.; May, S. *Applied Survival Analysis: Regression Modeling of Time-to-Event Data*. 2nd. Hoboken, NJ: John Wiley & Sons; 2008.
- Hurlburt MS, Garland AF, Nguyen K, Brookman-Frazee L. Child and family therapy process: Concordance of therapist and observational reports. *Administration and Policy in Mental Health and Mental Health Services*. 2010; 37:230–244.
- Kazdin AE. Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*. 2008; 63(3):146–159. [PubMed: 18377105]
- Kluger AN, DeNisi A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*. 1996; 119:254–284.
- Knap C, Koesters M, Schoefer D, Becker T, Puschner B. Effect of feedback of treatment outcome in specialist mental healthcare: meta-analysis. *British Journal of Psychiatry*. 2009; 195(1):15–22. [PubMed: 19567889]
- Lambert, MJ.; Barley, DE. Research summary on the therapeutic relationship and psychotherapy outcome. In: Norcross, JC., editor. *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients*. Oxford, England: Oxford University Press; 2002. p. 17–32.
- Lambert MJ, Harmon C, Slade K, Whipple JL, Hawkins EJ. Providing feedback to psychotherapists on their patients' progress: Clinical results and practice suggestions. *Journal of Clinical Psychology*. 2005; 61(2):165–174. [PubMed: 15609358]

- Lambert MJ, Whipple JL, Hawkins EJ, Vermeersch DA, Nielsen SL, Smart DW. Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology*. 2003; 10(3): 288–301.
- Levinson W, Lesser CS, Epstein RM. Developing physician communication skills for patient-centered care. *Health Affairs*. 2010; 29:1310–1318. [PubMed: 20606179]
- Makoul G. Essential elements of communication in medical encounters: The Kalamazoo consensus statement. *Academic Medicine*. 2001; 76:390–393. [PubMed: 11299158]
- Noell, GH.; Gansle, KA. The use of performance feedback to improve intervention implementation in schools. In: Hagermoser Sanetti, LM.; Kratochwill, TR., editors. *Treatment integrity: A foundation for evidence-based practice in applied psychology*. Washington, DC: American Psychological Association; 2014. p. 161-184.
- Riemer, M.; Bickman, L. Using program theory to link social psychology and program evaluation. In: Mark, MM.; Donaldson, SI.; Campbell, B., editors. *Social psychology and program/policy evaluation*. New York: Guilford; 2011.
- Riemer M, Douglas Kelley S, Casey S, Taylor Haynes K. Developing effective research-practice partnerships for creating a culture of evidence-based decision making. *Administration and Policy in Mental Health and Mental Health Services Research Special Issue: Research Practice Partnerships*. 2012; 39(4):248–257.
- Riemer M, Rosof-Williams J, Bickman L. Theories related to changing clinician practice. *Child and Adolescent Psychiatric Clinics of North America*. 2005; 14:241–254. viii. [PubMed: 15694784]
- Sapyta J, Riemer M, Bickman L. Feedback to clinicians: Theory, research and practice. *Journal of Clinical Psychology*. 2005; 61:145–153. [PubMed: 15609360]
- Shimokawa K, Lambert MJ, Smart D. Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*. 2010; 78:298–311. [PubMed: 20515206]
- Sparks JA, Kisler TS, Adams JF, Blumen DG. Teaching accountability: Using client feedback to train effective family therapists. *Journal of Marital and Family Therapy*. 2011; 37:452–467. [PubMed: 22007779]
- Twisk JWR, Smidt N, de Vente W. Applied analysis of recurrent events: A practical overview. *Journal of Epidemiological Community Health*. 2005; 59:706–710.
- Wampold, BE. *The great psychotherapy debate: Models, methods, and findings*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2001.
- Worthen VE, Lambert MJ. Outcome oriented supervision: Advantages of adding systematic client tracking to supportive consultations. *Counseling and Psychotherapy Research*. 2007; 7(1):48–53.
- Youn SJ, Kraus DR, Castonguay LG. The Treatment Outcome Package: Facilitating practice and clinically relevant research. *Psychotherapy*. 2012; 49:115–122. [PubMed: 22642519]

Figure 1. Kaplan-Meier Survival Plot for the SRF Topic Addressing Behavioral Issues by Randomized Group (feedback = 1, non-feedback = 0)

Note: The feedback group was coded as 1, with the non-feedback group coded as 0.

Figure 2.

Forest plot summarizing results of Cox regression for recurrent events where addressing content in sessions was predicted by feedback group, clinician- or client/caregiver-reported dynamic problem-specific severity, and their interaction.*

* This table only includes the simplest model chosen for interpretation for each topic by reporter (Cl = Clinician, CY = Client and/or caregiver). The results for all models including statistics are presented in the appendix (Table 1 for clinician-reported measures and Table 2 for client and/or caregiver-reported measures). All reported models were significant at an adjusted p -value of $p < 0.01$. S = Simple model (likelihood of addressing/ focusing on content was higher for clients whose clinicians were in the feedback condition compared to the non-feedback condition); A = Adjusted model (relative likelihood that feedback increased addressing/focusing on content after adjusting for cumulative problem alerts); I = Interaction model (the relative likelihood for the feedback condition of addressing/focusing on content where cumulative problem alerts were present). Where the interaction model was significant, the hazard ratio of feedback was calculated by summing the hazard ratios for feedback plus the interaction term where the cumulative problem alert = 1.

Table 1
Baseline characteristics of clients (N = 299) participating in the study

Characteristic	Control group (N = 149)		Feedback group (N = 150)	
	N	%	N	%
Gender				
Female	75	52%	68	46%
Male	68	48%	81	54%
Ethnicity				
Hispanic or Latino	11	10%	19	16%
Not Hispanic or Latino	104	90%	99	84%
Race				
White	83	67%	64	47%
Nonwhite	41	33%	71	53%
American Indian or Alaska Native	3	2%	1	1%
Black or African American	22	18%	37	27%
More than one race	13	11%	21	16%
Other	3	2%	12	9%
Age (mean $\pm$ SD years)	15.03 $\pm$ 1.82		14.52 $\pm$ 1.75	
Baseline SFSS (mean $\pm$ SD score)*				
Youth	2.33 $\pm$ 0.64		2.37 $\pm$ 0.71	
Caregiver	2.53 $\pm$ 0.74		2.59 $\pm$ 0.73	
Clinician	2.26 $\pm$ 0.56		2.41 $\pm$ 0.52	
Number of sessions (mean $\pm$ SD)	10.88 $\pm$ 9.70		10.15 $\pm$ 8.07	
Number of weeks in CFS (mean $\pm$ SD)	19.74 $\pm$ 15.35		16.93 $\pm$ 10.61	

* Possible scores for the SFSS range from 1 to 5, with higher scores indicating greater severity.

PTPB measure sources of session problem alerts by SRF topic and informant*

Table 2

SRF Topic	Informant ¹	Measure ²	Administration Schedule	# Problem Alert Items ³	Likert Scale for Items ³
Emotional issues	Y, Cg, CI	SFSS	Even weeks	13	1 – 5 (high)
Behavioral issues	Y, Cg, CI	SFSS	Even weeks	16	1 – 5 (high)
Family issues	Y, Cg, CI	SFSS	Even weeks	4	1 – 5 (high)
School/work issues	Y, Cg, CI	SFSS	Even weeks	3	1 – 5 (high)
Friends/peer issues	Y, Cg, CI	SFSS	Even weeks	6	1 – 5 (high)
Problems w/delinquent behavior	Y, Cg, CI	SFSS	Even weeks	3	1 – 5 (high)
Alcohol/substance use	Y, Cg, CI	SFSS	Even weeks	2	1 – 5 (high)
Harm to self or others	CI	SFSS	Even weeks	1	1 – 5 (high)
Client hope for future	Y	CHS-PTPB	Once/quarter	4	1 – 6 (low)
Client motivation for treatment	Y	MYTS	Twice/quarter	4	1 – 5 (low)
Therapeutic relationship w/client	Y	TAQS	Odd weeks	5	1 – 5 (low)
Client perceptions on counseling impact	Y	YCIS	Odd weeks	6	1 – 5 (low)
Client satisfaction w/services	Y	SSS	Once/quarter	4	1 – 4 (low)
Caregiver strain	Cg	CGSQ-SF7	Once/quarter	7	1 – 5 (high)
Caregiver satisfaction w/life	Cg	SWLS	Twice/quarter	5	1 – 7 (low)
Caregiver motivation for youth treatment	Cg	MYTS	Twice/quarter	7	1 – 5 (low)
Therapeutic relationship w/caregiver	Cg	TAQS	Odd weeks	5	1 – 5 (low)
Caregiver satisfaction w/services	Cg	SSS	Once/quarter	4	1 – 4 (low)

¹ Note: Y=Youth; Cg=Caregiver; CI=Clinician.

² Note: SFSS= *Symptoms and Functioning Severity Scale* (Cronbach's alpha = 0.92 – 0.94 depending on informant; Athay, Riemer, & Bickman, 2012); CHS-PTPB = *Children's Hope Questionnaire-Revised PTPB edition* (Cronbach's alpha = 0.87; Dew-Reeves, Athay, & Douglas Kelley, 2012); MYTS=*Motivation for Youth's Treatment Scale* (Cronbach's alpha = 0.82 – 0.84 depending on informant; Breda & Riemer, 2012); TAQS=*Therapeutic Alliance Quality Scale* (Cronbach's alpha = 0.85; Bickman, Vides de Andrade, Athay, Chen, DeNadai, Jordan-Arthur, & Karver, 2012); YCIS=*Youth Counseling Impact Scale* (Cronbach's alpha = 0.90; Kearns & Athay, 2012); SSS=*Service Satisfaction Scale* (Cronbach's alpha = 0.85-0.86 depending on informant; Athay & Bickman, 2012); CGSQ=*Caregiver Strain Questionnaire-SF7* (Cronbach's alpha = 0.89; Brannan, Athay, & Vides de Andrade, 2012); SWLS=*Satisfaction with Life Scale* (Cronbach's alpha = 0.89; Athay, 2012).

³ Problem alert are items rated at the top 25% of the distribution of item severity from the psychometric sample described in the PTPB manual (Bickman et al., 2010). For all measures, response choices were Likert-type items with varied scales. The directionality of the Likert scales for each measure is indicated by low (lower scores are more severe) or high (higher scores are more severe).

Table 3
Descriptive Information for All Clients (N = 299) on Session Content Addressed/Focused Upon and Cumulative Problem Alerts (Reported by Clinician or Combined Client &/or Caregiver) by SRF Topic

Session Report Form (SRF) Topic	Session content by topic			Clinician-reported Problem Alerts			Combined Client/caregiver Problem Alerts		
	N of clients with non-missing data	% Clients with any add/foc	% sessions with add/foc	N of clients with non-missing data	% Clients with any alert	Median # alerts	N of clients with non-missing data	% Clients with any alert	Median # alerts
Emotional issues	295	98%	82%	280	96%	17.0	269	97%	19.0
Behavioral issues	295	97%	82%	280	95%	15.5	269	96%	23.0
Family issues	295	97%	83%	280	83%	4.0	269	84%	6.0
School/work issues	295	94%	69%	280	85%	3.0	269	87%	5.0
Friends/peer issues	295	92%	66%	280	95%	8.0	269	93%	8.0
Problems w/delinquent behavior	291	52%	23%	280	79%	2.5	269	85%	3.0
Alcohol/substance use	290	51%	19%	280	42%	0.0	269	52%	1.0
Harm to self or others	293	58%	20%	280	45%	0.0			
Client hope for future	291	88%	58%				173	43%	0.0
Client motivation for treatment	293	84%	46%				121	41%	0.0
Therapeutic relationship w/client	293	67%	27%				268	60%	1.0
Client perceptions on counseling impact	292	79%	39%				132	61%	1.5
Client satisfaction w/services	293	73%	32%				184	23%	0.0
Caregiver strain	292	75%	38%				124	78%	3.0
Caregiver satisfaction w/life	292	63%	25%				151	54%	1.0
Caregiver motivation for youth treatment	291	69%	32%				97	38%	0.0
Therapeutic relationship w/caregiver	291	49%	17%				208	59%	1.0
Caregiver satisfaction w/services	290	58%	23%				130	5%	0.0