**Background**

While there is evidence that psychotherapy is effective for a majority of people, research consistently shows that some people do not achieve positive outcomes over the course of treatment (Finch, Lambert, & Schaalje, 2001). Studies have estimated that 30-50% of clients fail to respond to treatment, with 5-10% of clients actually worsening during treatment (Lambert & Ogles, 2004). Further, clinicians overestimate the change they are inducing in their clients (Walfish, McAlister, O’donnell, & Lambert, 2012), consistently failing to predict negative outcomes in their own clients and underestimating the incidence of negative outcomes in general (Hannan et al., 2005).

Routine outcome monitoring (ROM) aims to address this clinician blind spot by capturing client progress at regular intervals throughout treatment, ideally at every session, and providing that information back to the treating clinician in real-time. Building on this, patient-focused research uses ROM data from past clients to model average client change trajectories over the course of treatment, often stratified by initial distress and other client characteristics at intake. A client’s actual change during treatment is compared to their predicted change trajectory, and this information is then provided to clinicians as feedback to inform treatment (Boswell, Kraus, Miller, & Lambert, 2013). The feedback can alert a clinician when a client is not on track for a positive outcome (NOT) or inform the clinician that progress is on track (OT) with past similar clients. ROM has quickly become a recommended standard of care, with APA including ROM as part of effective evidence based care ﻿(American Psychological Association Presidential Task Force on Evidence-Based Practice, 2006).

* Introduce the present study and need for it- more research needs to be done

Many feedback systems have been developed: OQ-45 (Lambert, Kahler, et al., 2013), CORE-OM (citation), PCOMS ﻿(Duncan & Miller, 2008: Prescott et al., 2017), TOP (citation), COMPASS (citation). The methods behind the OQ-45 and the PCOMS will be reviewed here as examples of two of the most commonly used ROM instruments. The OQ-45 assesses client functioning across three domains: psychological symptoms, interpersonal problems, and social role functioning. It provides scores in each domain, as well as a total score, which is most commonly used.

In clinical practice, clinicians are presented with a graphical representation of their client’s scores over time, as well as a predicted change trajectory to which their client’s actual change can be compared. This predicted change trajectory is based on longitudinal treatment data from thousands of previous clients and is stratified by initial OQ-45 score. Change is modeled according to the dose-effect model, which characterizes change as a rapid initial decrease in symptoms, followed by increasingly more sessions needed to achieve the same amount of change. This is modeled as a decelerating logarithmic curve. The OQ-45 also provides several status alerts, indicating that clients are deviating significantly from the expected recovery curve in either the positive or negative direction. A positive alert can be indicative that a client is making progress more quickly than expected and may be ready to terminate therapy earlier, while a negative alert indicates that therapy is not progressing as quickly as expected and client might be at risk for treatment failure. Alerts are based on 80% tolerance intervals around the predicted scores at each session, with scores falling outside the upper limit of the tolerance interval indicating that they are at risk of being included in the 10% of clients likely to have a negative therapy outcome. The OQ-45 also provides clinical support tools and assessment for signal cases, a system for helping clinicians problem solve off track cases by identifying factors that have been shown to relate to client outcome.

The PCOMS assesses mental health functioning with the 4-item Outcome Rating Scale (ORS) and therapeutic alliance with the 4-item Session Rating Scale (SRS). It also uses expected trajectories of change for the ORS, based on Bayesian inference. ORS trajectories incorporate a client’s initial score, as well as their change at the current session relative to the initial score (Miller, Duncan, Sorrell, & Brown, 2005). Clients whose scores fall below the 50th percentile of these expected trajectories are identified as at risk (Anker, Duncan, & Sparks, 2009, p. 697), indicating that they are making less change than the average client. While other ROM instruments employ other methods to provide feedback on client progress, the OQ-45 and PCOMS represent two common examples.

* Review of literature on effectiveness of feedback
  + General positive effects of ROM
    - Although each ROM system uses somewhat different methods to derive alerts and provide feedback, feedback has been shown to improve outcomes and help prevent treatment failure across systems. A recently published meta-analysis including X studies using either the OQ-45 or PCOMS system found… (Lambert, Whipple, & Kleinstäuber, 2018)
    - (Shimokawa, Lambert, & Smart, 2010).
    - Ellsworth, Lambert, & Johnson, 2006; Lambert, Whipple, Bishop, et al., ﻿2002; Lutz et al., 2006; Spielmans, Masters, & Lambert, 2006
    - OQ (Harmon et al., 2007; Lambert & Shimokawa, 2011)
    - PCOMS (Duncan, 2012) (Anker, 2009)
    - Gondek 2016 review
    - Meta-analyses
      * Meta- analysis showing effect of *d* = .28 for all clients and *d*  = .53 for NOT clients (Shimokawa et al., 2010)
      * Fortney 2017 meta analysis
      * Lambert 2003 meta analysis
      * Lambert 2011 meta analysis
      * Lambert 2018 meta analysis
      * Østergård 2018 PCOMS meta analysis
  + Specific effects of feedback on different outcomes
    - OQ predicts deterioration in 85-100% of cases, with some false positives (Lambert et al., 2018)
  + Although there is a preponderance of research showing feedback to improve outcomes, there are some studies showing less promising results.
    - Meta-analysis showing small effect of *d* =.10 (Knaup, Koesters, Schoefer, Becker, & Puschner, 2009)
    - No overall effect, but effect for NOT clients using OQ (de Jong, van Sluis, Nugter, Heiser, & Spinhoven, 2012)
    - Kendrick 2016 meta analysis- used very strict inclusion criteria that eliminated studies with a stronger effect

With some exceptions outlined above, most research has shown a generally positive effect of feedback on client outcomes. Research is mixed, however, on whether this positive extends only to clients who go off track and are specifically at risk of a negative outcome, or whether clients already on track for a positive outcome also benefit from feedback. Generally, for clients already on track, receiving feedback that they are on track, or the absence of a NOT alert, has not been shown to improve outcomes (﻿Crits-Christoph et al., 2012; Harmon et al., 2007; de Jong et al., 2012; Lambert, Whipple, Smart, Vermeersch, & Nielsen, 2001; Lambert et al., 2002; Probst et al., 2013; Simon, Lambert, Harris, Busath), although some studies do find a positive effect of feedback for on track clients as well (Amble, Gude, Stubdal, Andersen, & Wampold, 2015). Some research shows that feedback may not improve outcomes for on track clients, but may shorten treatment, perhaps by indicating to therapists that the client has achieved the amount of change expected and may not need continued treatment (citation).

Beyond establishing that feedback improves outcomes in most studies, the literature is more mixed on how, when, and why feedback is effective, and answering these questions may shed light on the studies in which feedback was not found to improve outcomes

* How/why/when is feedback effective?
  + More broad research on feedback
    - Has to be actionable, provide new information, etc. See Lambert review.
  + Is it the alert that improves outcomes?
    - Trajectories of NOT patients were similar in FB and NFB conditions until the therapist in the FB condition was signaled that the patient was off track, but from that point forward the outcomes of patients in the FB condition improved significantly more than those in the NFB condition (Probst et al., 2013)
    - Same percentage of clients in feedback and no feedback conditions go off track (Amble, Gude, Stubdal, et al., 2015).
    - Clients in feedback and no feedback conditions who went off track had similar trajectories up to the point that the feedback condition clients received feedback, at which point their trajectory departed, indicating that the off track feedback was the effective component (Amble, Gude, Stubdal, et al., 2015)
    - Some research shows that slopes do change after a client receives an alert (Probst et al., 2013; Simon, Lambert, Harris, Busath, & Vazquez, 2012), but other research did not find an effect of feedback alert on slope (Amble, Gude, Ulvenes, Stubdal, & Wampold, 2015)
    - Although clients’ slopes improved after they received feedback, clients in a no feedback condition also had improved slopes after they went off track (when they would have received feedback), and the post-feedback slopes in the two conditions weren’t significantly different, failing to conclusively conclude that feedback results in improved rate of change. Instead, authors posit that signals tend to occur at high scores which are more likely to regress to the mean, potentially accounting for the decreasing slopes in both conditions (Amble, Gude, Ulvenes, et al., 2015)
    - Most ROM measures capture a single domain of distress (see TOP for an exception), and the literature is largely silent on how domain specific feedback and alerts compare to general feedback alerts. Further, it is unknown whether certain domains or areas of distress benefit from feedback more than others.

In addition to the scant research on mechanisms by which feedback improves outcomes, or fails to in some cases, there is similarly little research addressing the question of form whom feedback is most effective, and it is unclear whether certain types of clients benefit more from treatment that utilizes a feedback system.

* For whom is feedback effective?
  + One study found that feedback produced more pronounced effects for clients that started therapy with higher distressed, the who were most likely to otherwise have a poor outcome in therapy (Lambert et al., 2001). Another study, however, found that client initial distress did not moderate the effect of feedback (Amble, Gude, Ulvenes, et al., 2015).
  + Other prior research on client moderators of effectiveness of feedback?

These inconsistent findings in the extant literature suggest the need for additional research confirming the positive effect of feedback in actual practice and further exploring under what conditions and for whom it is effective. The present study aims to evaluate the impact of one specific feedback system developed for use in college counseling centers, as well as contribute to that body of literature by examining the effects of a multidimensional ROM feedback system in a large naturalistic data set.

**CCMH**

The Center for Collegiate Mental Health (CCMH) is a nationally representative practice research network (PRN) (Hayes, Locke, Castonguay, & Locke, 2011), built on a collaborative infrastructure involving multiple stakeholders, including university administrators, psychological researchers, industry partners, and over 600 university and college counseling centers. CCMH fills the primary goals of a PRN by facilitating the collection of information that will both inform clinical practice and advance research on the mental health services provided to UCC clients, while not adding substantial burden to everyday clinical practice.

The Counseling Center Assessment of Psychological Symptoms (CCAPS) is a routine outcome monitoring instrument developed by CCMH specifically for use in a college population. The CCAPS short form used here has 34 items capturing distress across seven domains: Depression (6 items), Generalized Anxiety (6 items), Social Anxiety (5 items), Academic Distress (4 items), Eating Concerns (3 items), Alcohol Use (4 items), and Hostility (6 items). In completing the CCAPS, clients are asked to rate themselves over the past two weeks on a Likert scale, from 0 (*not at all like me)* to 4 (*extremely like me*). Each subscale is scored by taking the average of the questions that load onto that subscale. As such, higher subscale scores indicate more distress, with scores ranging from 0 to 4. Although mean scores are used for research purposes, in clinical practice, CCAPS scores are displayed as percentiles, normed based on the nationally representative CCMH database of treatment seeking college students.

* Describe development of CCAPS feedback curves and methods
  + Previous ROM report provided CCAPS data in a tabular format, with each column representing a CCAPS subscale, and rows for each CCAPS administration. Although the CCAPS scores are used in research as the mean subscale score (ranging from 0-4), in clinical practice, subscale scores are presented as percentiles, normed on a distribution of clinical scores.
    - Figure of old report in appendix
  + Accomplishing one of the aims of a practice research network, CCMH used CCAPS data collected through clinical practice to build a feedback system to improve clinical outcomes. This feedback system introduced several new features into the CCAPS report, largely modeled after the OQ-45 feedback system. The first change was the addition of a graphical display of a client’s actual CCAPS scores on each subscale, allowing for ease of interpreting trends across administrations. Additionally, expected recovery trajectories based on past clients starting at a similar level of distress on that subscale were added alongside clients’ actual scores. Finally, an alert system was added to indicate if a client was off track from their expected recovery trajectory. This off-track alert is displayed as a blue dot, and no alert is displayed if a client’s scores are on track.
  + Beyond the actual changes to the CCAPS report, therapists who used the old and new report reported trickle down effects into how they conducted therapy, indicating that they brought the CCAPS into therapy to review with a client more because the report was more interpretable and the visual depiction of distress scores over time sparked conversation about how therapy was progressing and guided the focus of the session. The current research makes no attempt to isolate only the effects of a single component of the new CCAPS report (e.g. the effect off track alerts independent of the graphical depiction) or to isolate the effects of the new report from any impacts it may have had on how therapists actually conducted. Instead, this constellation of report changes are evaluated together and any behavioral changes are embraced as pathways by which feedback can have effects on therapy.
  + The methods behind the expected treatment trajectories and off-track alerts were based on those used for the OQ
    - Ns
    - Each subscale was independent
    - Binned by initial severity
    - Feedback alerts based on one-sided 90% tolerance intervals. No positive rapid response alerts are displayed, so only an upper limit to the tolerance interval was needed.
  + Limitation of clients that can’t alert because they start too high
* Development of the new feedback system exemplifies a practice research infrastructure, as CCAPS data collected as part of routine clinical practice was then used in research to better understand how clients change while in treatment, as well as used to develop the feedback tool that was implemented back into UCCs with the goal of informing everyday clinical practice and ultimately improving clinical outcomes. The present study will evaluate whether the data derived feedback system did positively impact client outcomes, and if so, for whom.

**Research questions**

* + The primary research question assesses the impact of the CCAPS feedback system. Did counseling center outcomes improve after the implementation of the new feedback system?
    - Because the CCAPS as a ROM measure was in place prior to the feedback system being implemented, any effects of the feedback system will be above and beyond the effect of ROM alone.
  + The second research question builds on prior research, which has largely been done in unidimensional measures of distress, to assess whether the effect of the new feedback system differed by subscale.
  + Is there a center effect for feedback?
    - The presence of a center effect would indicate the presence of center characteristics that moderate feedback’s effectiveness. Although such center characteristics are not measured in the present study, testing for a center effect will inform future research directions on identifying characteristics of centers for which feedback was especially helpful. Such characteristics may be able to be implemented elsewhere to improve positive effects of feedback.
  + Finally, several client moderators will then be evaluated to answer the question, for whom does feedback most improve outcomes?
    - Whether the client alerted or would have alerted
      * Answers question of whether this improves outcomes more for clients who alert or all clients
    - Mental health history items
      * Prior therapy
    - Frequency of CCAPS administrations
      * Centers have latitude to administer the CCAPS during treatment as frequently or infrequently as they chose. CCMH recommends that it is most effective when administered at every session, which many centers follow, while other centers administer at prespecified sessions (e.g. 1st, 3rd, and 7th). Is feedback more effective when the CCAPS is administered more frequently?
    - Total number of sessions
      * More effective for clients who have more sessions
    - Number of alerts a client receives
      * ?
    - Variability in client’s scores
      * Are alerts more effective for stable or unstable clients?
    - Whether the client was in other treatment modalities besides individual therapy

**Methods**

**Procedure**

Data for the present study were collected through the Center for Collegiate Mental Health (CCMH), which as mentioned above, is a practice research network of over 600 university and college counseling centers. Participating counseling centers collect data locally as part of clinical routine using standardized measures and can elect to contribute their center’s deidentified data to the centralized CCMH repository. Each year of CCMH data captures a single academic year, spanning from July 1 to June 30 of the following year. Four years of CCMH data from two discrete time periods are included in the present study, and centers are included if they contributed data to the repository during all four years. The first time period (2013-2015) captures data from before the updated CCAPS report was released, and the second time period (2016-2018) begins one year after the report was released on July 27, 2015. This intentional one-year gap accounts for the gradual adoption of the new report over the year following its release. While many centers began using the new CCAPS report as soon as it was made available, others began using it at a later date after having a chance to train staff in the new report. Dates on which centers upgraded to the new report within their electronic medical record system were consulted and any centers who had not updated by July 1, 2016 were excluded. This one-year gap also allows for centers to have time to become accustomed to the new report, ensuring that any clinical benefits coinciding with the release of the new report are not simply attributable to the report being something new. A dichotomous variable will be created indicating whether a client received treatment before or after the new CCAPS report and feedback system was implemented. These will be referred to as no feedback and feedback conditions. It is important to note that these two conditions occurred one after the other temporally, not simultaneously as in a randomized experiment.

**Measures**

* **CCAPS**
* **Standardized Data Set (SDS).** The SDS collects information on demographics, academics, and mental health history and is most often administered at the beginning of treatment (Hayes et al., 2011). The present study includes
  + SDS items included

**Analyses**

For each of the 8 CCAPS subscales, several client outcomes will be evaluated to determine whether outcomes improved in the feedback condition, answering research question one. For each outcome, results on each of the CCAPS subscales will be compared to determine whether the effect of feedback differed by subscale, answering research question two. To address research question three, a random effect for center will be included in each model, indicating whether the effect of the feedback system differed by center. In analyses for each outcome, only clients starting above the low cut point will be included, testing the impact of feedback on clients with some distress on each subscale.

The following outcomes will be evaluated for each subscale. First, the rate of deterioration, or reliable worsening, in each condition will be compared. Deterioration will be determined using the Jacobson and Truax method using the Reliable Change Index for each subscale (Jacobson & Truax, 1991). Due to ceiling effects on each subscale, some clients start with a score high enough that they are not able to deteriorate, and these clients will be removed from this analysis. This will also be modeled using two-level mixed effects logistic regressions for each CCAPS subscale. As feedback is specifically targeted at reducing deterioration in clients at risk for negative outcomes, it is hypothesized that clients in the feedback condition will be less likely to deteriorate.

Second, pre to post change will be compared. This will be modeled using a two-level mixed effects linear model, with clients nested within centers. Clients with a first and last CCAPS administration within 14 days respectively of their first and last individual appointments will be included. It is hypothesized that clients seen in the feedback condition will experience more change.

Finally, clients’ rate of change will be compared across feedback conditions to test whether clients’ symptoms across the CCAPS domains improve more quickly in the feedback condition. This will be modeled with a three-level mixed effects linear model, with sessions nested within clients nested within centers. In line with prior research, the effect of session will be log transformed (citation). Clients with at least two sessions will be included in these analyses, and only individual therapy sessions will be considered. It is hypothesized that clients seen in the feedback condition will have steeper slopes, experiencing more rapid change during treatment.

Although client outcomes are the focus of the present study, these are not the only outcomes that matter in evaluating such a system. Other outcomes that are left to future research include therapist satisfaction. Future directions for such research will be explored further in the discussion in connection with the findings from the present study.

Several moderators will be examined to determine for whom feedback is most effective, addressing research question four. Interactions between each moderator and the dichotomous feedback/no feedback condition variable will be added to each of the models outlined above. Interactions significant at the *p* < .001 level will be interpreted.

* Whether the client went off track during treatment
  + It is anticipated that clients in the feedback condition who received an off-track alert will have better outcomes than clients in the no feedback condition who went off track but did not receive the alert. This addresses the question of whether feedback improves outcomes more for clients who receive alerts. Literature is mixed.
* Initial CCAPS distress
  + Each client’s initial CCAPS score will be included to control for effects of baseline distress, and to test whether feedback is more beneficial for more highly distressed clients.
* Mental health history items
  + Prior therapy
* Frequency of CCAPS administration
  + Centers have latitude to administer the CCAPS during treatment as frequently or infrequently as they chose. CCMH recommends that it is most effective when administered at every session, which many centers follow, while other centers administer at prespecified sessions (e.g. 1st, 3rd, and 7th). It is anticipated that outcomes will improve as the CCAPS is administered more frequently, and that this effect will be stronger in the feedback condition.
* Total number of sessions
  + It is anticipated that clients with more sessions will have better outcomes, and that this effect will be stronger in the feedback condition.
* Variability in client’s scores
  + Is feedback more effective for stable or unstable clients?
* Whether the client was in other treatment modalities besides individual therapy
  + Although the CCAPS is typically only administered at individual therapy appointments, and not at other types of appointments like group or psychiatric…

**Results**

* Descriptives
  + % of clients who alerted on each subscale (or would have alerted) and % that alerted (or would have alerted) on at least one subscale
  + When in therapy alerts occur
  + Compare clients who alerted and didn’t
    - Average starting CCAPS
    - Average # of sessions

**Discussion**

* Important questions we weren’t able to answer
  + How often were therapists attending to the CCAPS and specifically to the feedback it provided? Does this mediate its effectiveness? We’d hope so!
  + Does the effect differ based on whether therapists discussed feedback, or CCAPS scores generally, with clients?
    - It’s not clear whether providing FB to both clients and therapists is more effective than providing to therapists alone
      * No difference (Shimokawa et al., 2010)
      * More effective when given to both (Knaup et al., 2009)
  + At a therapist level and structural center level, is the effectiveness of the new feedback system moderated by attitudes toward outcome monitoring or towards a change in routine? For example, a feeling that the change was imposed top down by center administrators could attenuate its effectiveness, potentially through therapists not attending to the feedback.
  + Is the effectiveness moderated by the way in which the new CCAPS system was introduced and if there was any training surrounding it?
  + Making sense of the mixed literature on whether FB is effective only for NOT clients using these therapist variables
    - It may be that the NOT signal itself is impactful if therapists only attend to feedback when a client alerts. This may be the case in systems where therapists carry higher caseloads. If therapists attend to feedback all throughout treatment, even if a client has not alerted (e.g. comparing a client’s visual trajectory to the expected trajectory, noting if a client is getting worse, even if not enough to alert), they may be more likely to see benefits from a feedback system all along, even in the absence of an alert. This fits with prior research showing that the effectiveness of feedback is moderated by therapists’ belief that FB is useful (de Jong et al., 2012). Further supporting the role of therapist attention to FB, the PCOMS, which explicitly instructs therapists to discuss progress and feedback with patients, has demonstrated positive effects across all patients, not just those who alert ﻿(e.g., Anker, Duncan, & Sparks, 2009; Reese, Norsworthy, & Rowlands, 2009; Reese, Toland, Slone, & Norsworthy, 2010), although this is not conclusive, as the PCOMS varies from other FB systems in other ways as well.
* Limitations
  + Not an RCT, and no true comparison group available
  + Can’t rule out the effect of time, which was perfectly confounded with the introduction of the profile report