PATIENT ACTIVATION IN PSYCHOTHERAPY INTERACTIONS: DEVELOPING
AND VALIDATING THE CONSULTATION INTERACTIONS CODING SCHEME

Sam Malins

University of Nottingham, UK

Nima Moghaddam

University of Lincoln, UK

Richard Morriss, Thomas Schröder, Paula Brown, Naomi Boycott, Chris Atha

University of Nottingham, UK

Correspondence concerning this article should be addressed to Sam Malins, NIHR

Collaboration for Leadership in Applied Health Research and Care - East Midlands,

Institute of Mental Health, University of Nottingham Innovation Park, Triumph Road,

Nottingham, NG7 2TU. Email: sam.malins@nottingham.ac.uk

Abstract

Objective. We describe the development of an instrument aiming to offer interaction-level feedback based on "patient activation": Client confidence and perceived ability to manage their health.

Method. Twenty-two session-transcripts from cognitive behavioral therapy with high-users of healthcare were analyzed thematically, producing themes describing in-session interactions. Themes were sub-categorized using patient activation theory into high and low activation presentations. Two coders new to the process were trained to use this Consultation Interactions Coding Scheme. Inter-rater reliability, convergent validity and clinical utility were assessed and illustrated with extreme cases.

Results. Good-to-excellent inter-rater reliability was achieved. The Consultation Interactions Coding Scheme, therapeutic alliance and therapist competence were correlated. Client engagement in session structuring interactions correlated with outcome. The highest Consultation Interactions Coding Scheme scorer showed multiple outcome-improvements, the lowest scorer reported deteriorations.

Conclusions. This study presents the Consultation Interactions Coding Scheme's psychometric properties and indicates the value of client engagement in session structuring.

Keywords: Patient activation, multimorbidity, Cognitive Behavioral Therapy, interaction style, frequent attendance

Word Count: 4,064

Patient Activation in Psychotherapy Interactions: Developing and Validating the

Consultation Interactions Coding Scheme

Patient activation is defined as the degree to which a person feels confident and able to be actively involved in managing their own health (Hibbard, Mahoney, Stockard, & Tusler, 2005). Patient activation has received little attention within psychotherapy research, despite widespread use in behavioral medicine. There are no measures of patient activation at the level of psychotherapy interaction and no studies have explored its ability to predict psychotherapy outcomes. This article presents the development and preliminary validation of the Consultation Interactions Coding Scheme, an instrument designed to rate psychotherapy interactions based on patient activation.

Patient activation is a significant and relatively new concept in health, which has become a well-used and -researched approach to improving self-management and health-behavior change (Hibbard and Gilburt, 2014). The Patient Activation Measure predicts a range of health outcomes and service use (Hibbard, et al., 2005). The Patient Activation Measure is also used as a health outcome in itself, because physical and mental health indicators improve when patient activation is targeted and increased (Hibbard and Gilburt, 2014). Improving patient activation is identified as an important way of supporting long-term health, following a recent large-scale trial showing that patient activation predicts future health and service use (NHS England, 2019).

Patient activation overlaps with related terms addressing patients' engagement with their healthcare and health services. These include health literacy, collaboration, shared decision making and engagement more generally. Patient activation is defined as a more specific concept than other engagement-related terms. It is focused on interactions between patient and healthcare professional, whereas engagement has more

broadly addressed overall health engagement (Graffigna, Barello, Bonanomi, & Lozza, 2015).

The Patient Activation Measure is a questionnaire measuring retrospective, self-reported attitudes, knowledge and behavior associated with patient activation. In practice, Patient Activation Measure scores and an algorithm-based interpretation are fed back to the assessing clinician with the aim of improving clinician responsiveness and effectiveness to support care tailored to the individual's need. This is similar to Feedback-Informed Treatment within psychotherapy, where feedback from progress-monitoring questionnaires is used to guide psychotherapists' responsiveness.

Yet, there is a recognized disconnection between Patient Activation Measure assessment and patient activation interventions (Armstrong et al., 2017): Most interventions designed to increase patient activation (and consequently improve self-management) focus on changing interaction-style within healthcare consultations (Deen, Lu, Rothstein, Santana, & Gold, 2011). However, the Patient Activation Measure does not directly assess interaction behaviors in healthcare consultations or give feedback of sufficient detail to inform these behaviors, despite the evident importance of interaction-style to improving patient activation.

The problems identified with the Patient Activation Measure are mirrored in psychotherapy's Feedback-Informed Treatment: Both the Patient Activation Measure and Feedback-Informed Treatment could benefit from greater detail on in-session processes. Current best practice in Feedback-Informed Treatment does not typically extend to interaction-level feedback, despite its apparent importance. This is likely to mean that specific interaction-level micro-processes requiring corrective practice are harder to identify and feed back to therapists, inhibiting development of therapist effectiveness (Goldberg et al., 2016). Taken together, current evidence suggests that an

observational, interaction-focused rating system could enhance the deliberate practice method and assessment of patient activation.

This paper reports the development of the Consultation Interactions Coding Scheme, which aims to assess whether patient activation can be accurately rated at specific in-session interactions and whether these ratings correlate with outcomes. This detailed approach accounts for the same type of interaction being viewed as helpful or unhelpful at different points in a session (Swift, Tompkins, & Parkin, 2017), which is not addressed by more typical total behavior counts or global scores. Patient activation has been assessed within mental health settings (e.g. Green et al., 2010) and interactions have been rated for activation within diabetes care (Williams et al., 2005). However, interaction-level analysis of patient activation has not been conducted in psychotherapy.

This paper aimed to establish the inter-rater reliability, convergent validity and clinical utility of the Consultation Interactions Coding Scheme amongst a sample of high healthcare utilizers with multimorbidity. This client group was selected because of the likely impact patient activation would have, as multimorbid health problems are heavily reliant on patients' self-management skills (Hibbard and Gilburt, 2014).

Method

Design

To develop and test the Consultation Interactions Coding Scheme, we applied a sequential exploratory design (Creswell, Plano Clark, Gutmann, & Hanson, 2003). This is a two-phased mixed-methods design: beginning with qualitative data analysis to explore a phenomenon (in this case, themes of interaction within CBT sessions) and developing an instrument (gauging levels of patient activation within each theme presentation) which can be quantitatively tested in the second phase.

Participants

Thirty-two of 87 participants assessed (37%) consented to receive Cognitive Behavioral Therapy (CBT) in a case series (Malins et al., 2016). Eleven of the 32 (34%) consented to session data being anonymously reported and are included in this study (Table 1). Case series participants were approached after being identified from electronic records as having primary care consultation rates above the 90th centile for the previous two years, typically related to multimorbidity (average five chronic health conditions). For each participant, a true random number generator (https://www.random.org) was used to select one session-recording from the earlier half of therapy, and one from the latter half. Recordings were transcribed verbatim, disguising identifiable information.

 [Table 1]

Therapists

Two psychotherapists provided CBT to participants, one male and one female with doctoral and masters' level training respectively. Therapists were not blind to the purpose of the study and were not trained in improving patient activation.

Thematic analysis of CBT sessions

A two-stage thematic analysis was conducted using an adapted version of Braun and Clarke's (2006) method. In the first stage, each of two researchers (SM and NM) took half of the sample (five participants each with one shared) and completed a descriptive, inductive thematic analysis on selected session transcripts. This open approach aimed to identify "what is talked about in this session and how?" Having each looked at different parts of the dataset, similarities and differences in emerging themes were used to develop a theme template. A joint thematic model of interaction themes was refined through

discussion and review. Thematic analysis continued until a set of superordinate themes were established which were each represented in at least half of the transcripts analyzed, with coherence amongst themes.

In a second stage, theory-informed, outcome-orientated, deductive analyses of established themes were carried out. Whilst blind to outcomes, the two researchers reanalyzed and grouped transcript data within each super-ordinate theme into three types of theme presentation: (1) Those deemed high in patient activation (2) Those deemed low in patient activation and (3) Those where no observable distinction could be made. Patient activation theory was used to decide which types of interaction belonged to each group. In particular, patient activation theory suggests that high patient activation is expressed when someone is more engaged and confident in their ability to manage their health (Hibbard and Gilburt, 2014).

Quality assurance methods. Reflective journals were kept by researchers to help them consider how existing knowledge and beliefs might influence their reading of transcript data. Using these journals helped researchers set aside biases and engage with the data more openly. Multiple coding added quality assurance to the themes developed, as similar forms of major themes were present in both researchers' analyses. Researchers maintained an auditable process to ensure that emerging themes were grounded in and traceable to source data. These methods were congruent with the researchers' critical realist stance, which recognizes that personal interpretations can make it harder to get close to phenomena of interest, but effective research strategies can help (Fairclough, Jessop, & Sayer, 2002).

The results of the thematic analysis and coding methods were aggregated into a manual named the Consultation Interactions Coding Scheme (Malins et al., 2018). The Consultation Interactions Coding Scheme provides clear descriptors for rating patient

activation within each interaction, rather than global session scores alone. Using the same sample of CBT sessions, assessment of reliability, validity and an extreme case illustration were conducted.

Inter-rater reliability (IRR)

The two original researchers independently applied the initially developed Consultation Interactions Coding Scheme to the same 12 transcripts in batches of four with discussion of coding reliability between each batch. Comparison of the two researchers' ratings was used to progressively refine the Consultation Interactions Coding Scheme. Coding manual refinement was completed once a criterion-level intra-class correlation coefficient (ICC) of .70 had been achieved for each theme under a two-way, random effects model.

Two new researchers (PB and NB) blind to outcome, who had not been involved in the Consultation Interactions Coding Scheme's development, were trained to use it. Training involved three two-hour training sessions, using segments of six transcripts with two hours of coding completed between sessions. After training, feedback was sought from new coders, including reflections on ambiguity in coding definitions and overlap between themes. The coding scheme was amended, accounting for feedback. All four researchers were then given a final version of the Consultation Interactions Coding Scheme and the 10 remaining transcripts to assess IRR.

Convergent Validity

Each of the two randomly selected sessions per client-participant were rated by at least two of the four researchers using the Consultation Interactions Coding Scheme and a six-item observer-adapted working alliance inventory (WAI-O; Falkenström, Hatcher, Skjulsvik, Larsson, & Holmqvist, 2015; Horvath and Greenberg, 1989). Each session was also assessed by a different single independent rater (CA) using the

cognitive therapy rating scale – revised (CTS-R; Blackburn et al., 2001). The WAI-O and the CTS-R are commonly used global assessments of psychotherapy interactions focused on the therapeutic relationship and CBT competence respectively. Assessing correlations between globally rated interaction qualities on these measures and the Consultation Interactions Coding Scheme aimed to evaluate convergent validity.

Predictive Validity

The Consultation Interactions Coding Scheme, WAI-O and CTS-R scores for earlier and later sessions were separately correlated with clinical outcome scores between baseline and six-month follow-up for initial assessment of predictive validity. All outcomes were collected in person by researchers independent of psychotherapy provision at baseline and six-month follow-up. Outcomes included:

- The Acceptance and Action Questionnaire II (AAQ II), a measure of psychological flexibility (Bond et al., 2011).

- The seven item Generalized Anxiety Disorder scale (GAD-7), a measure of generalized anxiety symptoms (Spitzer, Kroenke, & Williams, 2006).

- The 18-item Health Anxiety Inventory (HAI), a measure of health anxiety (Salkovskis, Rimes, Warwick, & Clark, 2002).

- The EQ-5D-3L, a measure of quality of life (EuroQol Group, 1990).

- The 15-item Patient Health Questionnaire (PHQ-15), a measure of somatic symptoms (Kroenke, Spitzer, & Williams, 2002).

- The nine item Patient Health Questionnaire (PHQ-9), a measure of depression symptoms (Kroenke, Spitzer, & Williams, 2001).

- The Ruminative Responses Scale (RRS), measuring rumination as a response when feeling sad or depressed (Treynor, Gonzalez, & Nolen-Hoeksema, 2003).

- The Short-Form 36, a measure of general mental and physical health (mental component scale: MCS; physical component scale: PCS; Ware, Kosinski, & Gandek, 2005).

Extreme case illustration. To illustrate the Consultation Interactions Coding Scheme's discriminative validity, the two participants with the highest and lowest scores on the Consultation Interactions Coding Scheme were compared on clinical outcome, WAI-O and CTS-R scores.

Statistical Analysis

For each Consultation Interactions Coding Scheme theme, IRR assessment was based on four independent ratings of each theme at each turn-of-speech, meaning that data units were defined as transcribed turns-of-speech and content units were Consultation Interactions Coding Scheme themes rated by patient activation level (-2 to +2). If neutral assessments of patient activation were made or themes were rated as absent, these were scored as 0. Transformed scores were calculated by multiplying the percentage of transcribed words coded at each level of each theme by the score for the given level. For example, if 10% of a transcript was coded as Action Planning and Idea generation (API) at the +2 level, this would contribute a score of 20 to the total for the API theme (+2 x 10). This calculation was carried out for all themes on all transcripts, and the total theme scores were then compared across raters. The exception to this rule was the problem and context description theme (*PCD*). The *PCD* score was the percentage of the transcript coded *PCD*. Assessment of IRR was also conducted at the level of individual speech turns. Therefore, additional ICCs were calculated on an interaction turn-by-turn basis.

Two-way, random effects, consistency, average-measures ICCs were used to assess IRR with SPSS 24. Binary turn-by-turn ratings of the *PCD* theme were assessed

using Siegel and Castellan's (1988) kappa, averaged across each coding pair. Each session was represented by one summary score for each Consultation Interactions Coding Scheme theme by averaging scores between raters. Correlations among process and outcome scores were calculated using two-tailed Spearman's rho correlations.

To elaborate the clinical significance of extreme case comparisons, changes that cross published clinical cut-off boundaries and/or meet published criteria for minimal clinically important change were of particular interest. Where published cut-offs were not available, Jacobson and Truax's (1991) criteria were used to compute clinical cut-offs and the Reliable Change Index (RCI), for minimal clinically important change.

Results

Inductive Thematic Analysis

From the initial inductive thematic analysis, nine interaction themes were developed. These themes were organized into an interaction process model, described below (Figure 1). In terms of interaction contents, the CBT conversation process was seen to typically begin with descriptions of problems or their context (*Problem or context description: PCD),* in which therapist and client delineated the nature of a difficulty, the context in which it occurred, and/or its impact. If session discussions moved on from *PCD*, an analysis of problems often ensued *(Problem analysis and understanding: PAU)*, sometimes mediated by relevant *Information discussion (IDI)*. Identification of potential solutions or related information could lead the conversation into action plans for change or ideas about what could be done to improve the situation *(Action planning and idea generation: API).* Feedback from action plans and ideas was sometimes given, typically after they had been tested out in the session or between sessions. Feedback included *Evaluations of the self or therapy (EST)* alongside reports of change or the absence of change from therapeutic interventions *(Noticing change or*

otherwise: NCO). Feedback from action plans also informed the next cycle of the conversation process – for example the changes noticed could lead to a conversation about a new understanding reached *(PAU)* and consequent action plans *(API).* Conversely, if no change was observed, continuous description of related problems could ensue *(PCD)* or an analysis of the failure might take place *(PAU).* The joint structuring, choice and focusing of these movements between themes was identified as a separate anchoring theme for all other themes *(Structuring and task focus: STF).*

In global terms, the degree to which session conversations integrated and moved naturally between themes seemed important to the overall flow *(Integration of themes: IOT).* The mutuality of task and idea development between therapist and client also appeared to have an important impact on theme expressions *(Collaborative flow: COF).*

 [Figure 1]

Deductive Thematic Analysis

Each super-ordinate theme was sub-categorized into extracts where there were observable, verbal indications of engagement with, confidence in and/or facilitation of health improvement. These interactions were categorized positively (+1 or +2). Conversely, where there were observable, verbal indications of reluctance to engage and/or low confidence about health management, these were categorized negatively (-1 or -2). The criteria for these classifications were developed into a rating scheme where higher scores indicated higher patient activation and vice versa. Ratings were assigned to one of five ordered levels, ranging in valence and magnitude: from -2 for interactions deemed very low in patient activation, to +2 indicating interactions deemed very high in patient activation with a midpoint of 0 denoting a neutral or undiscernible level of patient activation. Level descriptors were characterized in a similar way across themes (exemplars for all themes are presented in supplementary materials).

Global themes *IOT* and *COF* applied an overall score using the above scale across each session as a whole. Problem description interactions (*PCD*) were deemed a negative outcome predictor, so the percentage of the session coded *PCD* was calculated rather than rated on a scale. This was because large amounts of problem description would give less opportunity for interactions where high patient activation could be generated or expressed.

Inter-rater reliability. Analysis included ratings of 3,134 turns of speech from 10 CBT sessions with five clients and two therapists, rated by four independent raters. Only the 10 transcripts that had not been rated in Consultation Interactions Coding Scheme development were used to assess IRR. All transformed theme scores achieved ICCs within the "excellent" range of agreement (ICCs = .84 to .97), except Information Discussion (*IDI*; ICC = .60) which fell within the "good" range (Cicchetti, 1994). All turn-by-turn theme ratings achieved ICCs within the good-to-excellent range, (ICCs = .69 to .80; Cicchetti, 1994). Turn-by-turn, the problem and context description theme (*PCD*, rated present/absent) showed moderate to substantial agreement ($\kappa = .54$ to .61).

Convergent Validity

Correlations across all 22 rated sessions showed significant, very strong Spearman's rho correlations between Consultation Interactions Coding Scheme themes and the WAI-O (r_ss = .72 to .91; $p < .001$), with the exceptions of structuring and task focus (*STF*; $r_s = .36$; $p = .102$) and information discussion (*IDI*; $r_s = .08$; $p = .719$). Significant, moderate-to-strong correlations were found between five of the nine Consultation Interactions Coding Scheme themes and CTS-R scores (Table 2). As predicted, the WAI-O and CTS-R had significant negative correlations with the problem and context description theme (WAI-O: $r_s = -.71$, $p < .001$; CTS-R: $r_s = -.47$, $p = .029$).

[Table 2]

Predictive Validity

The valence of all correlations is adjusted so that positive correlations consistently indicate improvement in clinical outcome. Significant, strong Spearman's rho correlations were found between structuring and task-focus (*STF*) interactions in the earlier half of therapy and improvements in: depressive symptoms (PHQ-9: $r_s = .63$; $p = .049$), rumination (RRS: $r_s = .67$; $p = .036$), and psychological flexibility (AAQ: $r_s = .65$; $p = .043$). Non-significant, moderate correlations were observed with improvements in generalized anxiety (GAD-7: $r_s = .59$; $p = .056$), health anxiety (HAI: $r_s = .49$; $p = .186$), and the mental component scale of the SF-36 ($r_s = .43$; $p = .190$). No significant outcome correlations were observed with the WAI-O, the CTS-R, or other Consultation Interactions Coding Scheme themes for earlier or later sessions.

Outcomes from Extreme Case Examples[1]

Descriptions of the highest and the lowest Consultation Interactions Coding Scheme scorers are given with a summary of clinical outcomes and compared to Consultation Interactions Coding Scheme, WAI-O and CTS-R scores.

Highest scorer. Participant 2 was a 35-year-old mother of two who periodically suffered with anxiety, depression, and bulimia. Outcomes indicated reliable improvement in generalized anxiety (GAD-7, 20-10), health anxiety (HAI, 23-17), depressive symptoms (PHQ-9, 16-10), psychological flexibility (AAQ-II, 44-33), rumination (RRS, 71-55) and the SF-36 (MCS, 19.8–26.9; PCS, 47.2–58.4), alongside small deterioration in quality of life that was not clinically significant (EQ-5D, .85–.81).

---

[1] Demographic details of both cases are similar but disguised to protect anonymity.

Lowest scorer. Participant 9 was a 78-year-old retired businessman who had developed multiple health problems over time. Reliable improvement in quality of life was observed (EQ-5D, .16–.62), but also reliable deterioration in psychological flexibility (AAQ, 7–31), depression (PHQ-9, 3–10), generalized anxiety (GAD-7, 0–6), and health anxiety (HAI, 29–44), with no change in any other outcomes.

Contrasting clinical outcomes were accompanied by Consultation Interactions Coding Scheme scores divergent in polarity (Consultation Interactions Coding Scheme content average: 13.33 vs. -3.61; global average: 1.31 vs. -1.06). The biggest contrast was in the structuring and task-focus theme (*STF*, 8.55 vs. -28.7). Participant 2 and 9's averaged WAI-O (5.8 vs. 3.1) and CTS-R (45 vs. 29) scores showed similar disparities.

Discussion

This study indicates that individual psychotherapy interactions can be reliably rated based on patient activation and that scores are associated with therapeutic alliance and therapist competence. There may also be an important relationship between specific turn-by-turn interactions and clinical outcomes, particularly patient activation in structuring and task-focus interactions.

New raters were trained to use the Consultation Interactions Coding Scheme with good-to-excellent inter-rater reliability. Most Consultation Interactions Coding Scheme themes showed correlation with therapeutic alliance and therapist competence indicating good convergent validity and potential for clinical utility of the Consultation Interactions Coding Scheme. Importantly, therapeutic alliance and therapist competence were both negatively correlated with problem description interactions. This indicates that more time focused on problem description may be associated with reduced therapeutic alliance and therapeutic competence amongst people with multimorbidity.

A noteworthy finding is that structuring and task-focus interactions in early

sessions correlated with multiple health outcomes, whilst no other interaction themes or process measures did. This suggests that actively engaging clients in the process of structuring sessions and therapeutic task choice early in therapy may be important to overall outcome, particularly if clients are experiencing complex health needs. The contrast in outcomes between the highest and lowest Consultation Interactions Coding Scheme scorers supports the proposition that turn-by-turn rated interactions may be associated with outcome. The large contrast in patient activation within interactions about structuring and task choice adds further support to the importance of these interactions for people with multimorbidity.

Typically, interaction rating schemes lead to total behavioral counts or global scores on identified dimensions. Turn-by-turn analysis establishes the specificity of raters' judgements for each unit of measurement. This provides a stronger sense of how well each rater agrees on specified data points. Otherwise, spurious chance agreement on overall scores could not be ruled out. Achieving reliable turn-by-turn analysis also allows easier identification of clinically important changes that occur within the session. Such changes can affect outcome and may not be picked up by total score analyses (Swift, et al., 2017). Overall, results suggest that a coding manual and training method has been developed with sufficient clarity, specificity, and reliability to provide a foundation upon which further study of interaction-level patient activation can be built.

This study was limited by low IRR for some themes assessed. Adequate IRR was not achieved for information discussion interactions, which may relate to the comparatively small number of interactions to which the theme was applied. It is also unclear why structuring and task focused interactions did not achieve convergent validity with therapeutic alliance or therapist competence. It may indicate that structuring interactions is a considerably different construct to other interaction types,

given the potential for predictive validity. Alternatively, it may be that this was a chance finding, due to multiple comparisons made with several Consultation Interactions Coding Scheme themes.

Given that all interaction themes were rated similarly, it is notable that no other Consultation Interactions Coding Scheme theme correlated with any outcome. As with convergent validity, this may highlight structuring interactions as particularly important, but it may alternatively suggest that some themes are redundant for outcome prediction. Larger-scale assessment would clarify whether low power affected the current results. Furthermore, clarity is required to confirm how well the construct of patient activation is being measured by the Consultation Interactions Coding Scheme. The established Patient Activation Measure was not included in this study for concurrent validity, because patient activation was not the primary focus of the source project.

Reliable turn-by-turn Consultation Interactions Coding Scheme rating provides the potential to offer greater specificity in guidance for individual therapists: Identifying the interactional skills most important for outcome. The Consultation Interactions Coding Scheme also provides a complementary method for assessing patient activation to global questionnaire-based assessment by providing specific information on how clinical interaction-style could be used to improve patient activation in healthcare consultations. In particular, reported findings suggest psychotherapists should facilitate client engagement in session-structuring and task choice early in therapy.

The high level of consistency in scoring suggests the Consultation Interactions Coding Scheme will be sufficiently dependable for future outcome-prediction studies. Therefore, larger-scale assessment of the Consultation Interactions Coding Scheme's predictive validity within a similar population would clarify whether low power affected the current results. Incorporating concurrent Patient Activation Measure and

Consultation Interactions Coding Scheme assessment would provide comparison with the gold-standard patient activation assessment. The Consultation Interactions Coding Scheme has potential applicability amongst other clinical groups and psychotherapies, which should be assessed in future research.

This study suggests that patient activation can be reliably rated at an interaction level in CBT for clients with complex problems. Furthermore, in-session patient activation is related to therapeutic alliance and therapist competence and may predict outcomes. Specifically, greater client engagement in session-structuring during early sessions may predict clinical outcomes.

References

Armstrong, N., Tarrant, C., Martin, G., Manktelow, B., Brewster, L., & Chew, S. (2017). *Independent evaluation of the feasibility of using the Patient Activation Measure in the NHS in England-Final report* London: The Health Foundation.

Blackburn, I., James, I. A., Milne, D. L., Baker, C., Standart, S., Garland, A., & Reichelt, F. K. (2001). The revised cognitive therapy scale (CTS-R): psychometric properties. *Behavioural and cognitive psychotherapy, 29*(04), pp. 431-446.

Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., . . . Zettle, R. D. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire–II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior therapy, 42*(4), pp. 676-688.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), pp. 77-101. doi:10.1191/1478088706qp063oa

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment, 6*(4), p 284.

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003).

Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie

(Eds.), *Handbook on mixed methods in the behavioral and social sciences* (pp.

209-240). Thousand Oaks, CA: Sage Publications.

Deen, D., Lu, W.-H., Rothstein, D., Santana, L., & Gold, M. R. (2011). Asking

questions: the effect of a brief intervention in community health centers on

patient activation. *Patient education and counseling, 84*(2), pp. 257-260.

EuroQol Group. (1990). EuroQol - a new facility for the measurement of health-related

quality of life. *Health Policy, 16*, pp. 199-208.

Fairclough, N., Jessop, B., & Sayer, A. (2002). Critical realism and semiosis. *Alethia,

5*(1), pp. 2-10.

Falkenström, F., Hatcher, R. L., Skjulsvik, T., Larsson, M. H., & Holmqvist, R. (2015).

Development and validation of a 6-item working alliance questionnaire for

repeated administrations during psychotherapy. *Psychological assessment,

27*(1), p 169.

Goldberg, S. B., Babins-Wagner, R., Rousmaniere, T., Berzins, S., Hoyt, W. T.,

Whipple, J. L., . . . Wampold, B. E. (2016). Creating a climate for therapist

improvement: A case study of an agency focused on outcomes and deliberate

practice. *Psychotherapy, 53*(3), p 367.

Graffigna, G., Barello, S., Bonanomi, A., & Lozza, E. (2015). Measuring patient

engagement: development and psychometric properties of the Patient Health

Engagement (PHE) scale. *Frontiers in psychology, 6*, p 274.

Green, C. A., Perrin, N. A., Polen, M. R., Leo, M. C., Hibbard, J. H., & Tusler, M.

(2010). Development of the patient activation measure for mental health.

Administration and Policy in Mental Health and Mental Health Services Research, 37(4), pp. 327-333.

Hibbard, J., & Gilburt, H. (2014). *Supporting people to manage their health: An introduction to patient activation.* London, UK: The King's Fund.

Hibbard, J., Mahoney, E. R., Stockard, J., & Tusler, M. (2005). Development and testing of a short form of the patient activation measure. *Health Services Research, 40*(6p1), pp. 1918-1930.

Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of counseling psychology, 36*(2), p 223.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of consulting and clinical psychology, 59*(1), pp. 12-19.

Kroenke, K., Spitzer, R., & Williams, J. (2001). The PHQ-9: the validity of a brief depression severity measure. *Journal of Internal Medicine, 16*, pp. 606-613.

Kroenke, K., Spitzer, R., & Williams, J. (2002). The PHQ-15: the validity of a new measure for evaluating the severity of somatic symptoms. *Psychosomatic Medicine, 64*, pp. 258-260.

Malins, S., Kai, J., Atha, C., Avery, A., Guo, B., James, M., . . . Morriss, R. (2016). Cognitive behaviour therapy for long-term frequent attenders in primary care: a feasibility case series and treatment development study. *British Journal of General Practice, 66*(651), pp. e729-736.

Malins, S., Moghaddam, N., Morriss, R., Schroder, T., Cope, N., & Brown, P. (2018). Consultation Interaction Coding Scheme 1.6. Retrieved 6 Nov, 2018 from https://figshare.com/articles/Consultation_Interaction_Coding_Scheme_CICS_/7302386.

NHS England. (2019). Supported self-management and patient activation. Retrieved 18

    June, 2019 from https://www.england.nhs.uk/ourwork/patient-participation/self-

    care/patient-activation/.

Salkovskis, P., Rimes, K., Warwick, H., & Clark, D. (2002). The Health Anxiety

    Inventory: development and validation of scales for the measurement of health

    anxiety and hypochondriasis. *Psychological Medicine, 321*, pp. 843-853.

Siegel, S., & Castellan, N. (1988). *Nonparametric statistics for the behavioral sciences*

    New York, NY: McGraw-Hill.

Spitzer, R., Kroenke, K., & Williams, J. M. G. (2006). A brief measure for assessing

    generalised anxiety disorder; the GAD-7. *Archives of Internal Medicine, 146*,

    pp. 1092-1097.

Swift, J. K., Tompkins, K. A., & Parkin, S. R. (2017). Understanding the client's

    perspective of helpful and hindering events in psychotherapy sessions: A micro-

    process approach. *Journal of Clinical Psychology, 73*(11), pp. 1543-1555.

Treynor, W., Gonzalez, R., & Nolen-Hoeksema, S. (2003). Rumination reconsidered: A

    psychometric analysis. *Cognitive Therapy and Research, 27*(3), pp. 247-259.

Ware, J. E., Kosinski, M. A., & Gandek, B. (2005). *SF-36 Health Survey: Manual and

    interpretation guide* Lincoln: Quality Metric Inc.

Williams, G. C., McGregor, H., Zeldman, A., Freedman, Z. R., Deci, E. L., & Elder, D.

    (2005). Promoting glycemic control through diabetes self-management:

    evaluating a patient activation intervention. *Patient education and counseling,

    56*(1), pp. 28-34.

Supplementary Table. *Exemplar quotes for all levels of each theme of the Consultation Interaction coding scheme*

Theme	Level	Exemplar Quote
Action planning and idea generation (**API**)	+2	**T**: is it [ ], like, this is it, you know, this is just life, it's always like this? **C:** No, I ought to make more of an effort to do my exercises and look after my back. I have been looking in my pain management book recently and thinking, yes I must do that... [Participant 5]
	+1	**T**: Yeah, what we did. So kind of like, we did about 12 minutes, so maybe a little bit more than what we just did, if you can do that each day. Like a gradual building process. **C**: If I get up 10 minutes early and do it in the morning, it could set me up for the day really couldn't it? [Participant 4]
	0	**T**: There's 2 key bits to it. One about, kind of, today and tomorrow, planning ahead what's going to make life easier, what's going to be kind of pacing it for the weekend. **C:** OK, right. [Participant 11]
	-1	**T**: So is this all about keeping mobile? **C**: I can't walk far so I'll have to have a taxi. Can't do it. [Participant 8]
	-2	**T**: Do you ever use kind of warmth or anything? What did the pain clinic say? **C**: This is another thing. I just cannot get warm. I sent for (Dr Smith), it was last year and I says my feet and my legs are frozen [Participant 8]
Evaluations of self or therapy (**EST**)	+2	**C**: My initial feeling was I'm pleased with myself for being honest with myself [ ] you know feeling, this is ironic! Confident enough to be honest with them about how I'm feeling. **T**: Why is that ironic? [Participant 2]
	+1	**T**: Mmm, how [has] it been particularly off the back of last time? **C**: I mean, this has been very helpful, very useful, it's been good. [Participant 5]
	0	**T**: you have been brave [ ] those are really big steps and really valuable steps [ ] and that is what you're telling yourself every time you [ ] let water go on your face a little bit more [ ]. Did you get round to kind of doing [a diary]? **C:** Well, I didn't... [Participant 7]
	-1	**C:** Back to square one aren't we. **T**: Is that what it feels like to you? [Participant 6]
	-2	**C**: I just don't think I've done anything right in life, I really don't. I don't see my point being here at all. And I keep going to see [my GP] and I feel like a hypochondriac. [ ] I just feel like nobody needs me. **T**: Yeah. [Participant 4]

Information Discussion (IDI)	+2	C: It would be really helpful if we could talk about how panic attacks work like you mentioned last time, so I know what I'm dealing with. T: Yeah, great idea. Shall we look at a recent example? *[Hypothesised example]*
	+1	T: What you think about, how you feel physically and emotionally [ ] and what you do are all linked [ ] so for example if you kept on making time for yourself [ ] over time you will feel less guilty, [ ] you will have less of those thoughts of feeling selfish [ ]. C: I need to practise that then, that's the first step then… [Participant 3]
	0	T: In a kind of very basic formulation that we do in cognitive behavioural therapy, so it can be thoughts about the situation, OK, physiology is about body sensations, how you feel, behaviour basically is what you do, OK, and emotion is how you feel. [ ] Is that OK? C: Uh-huh. [Participant 1]
	-1	T: You can also get headaches with dehydration. So it might be possible that you know with your - C: That would be a miracle then. [Participant 9]
	-2	T: But it is still good to drink good amounts, 'cause it's good for your kidneys, it's good for your brain. C: I haven't got a brain so I. [Laughter] My eyes are back to normal. I see them again, soon. We've got it all sorted out now. You know that funny letter I wrote... [Participant 9]
Noticing change or otherwise (NCO)	+2	C: I have noticed this week actually that when I've been ill and been coughing a lot and not been able to breathe I got very anxious. And I thought well there's no need to do this, you know it's not life threatening, it's nothing really bad. And I sort of talk myself down again. T: Good, good. More self-aware about it. [Participant 10]
	+1	T: Yeah, and you are saying like there is a drop in the pain actually after [the graded pain killer medication reduction]? C: Yes. As I say it did improve. [Participant 6]
	0	T: Yeah, so [your sense of achievement] was feeling different then anyway. C: Yeah. [Participant 11]
	-1	T: Well, aside from that how have things been? C: Oh, not too good I don't think in the fact that I have been in a lot of pain, I mean the worst it's ever been [Participant 7]
	-2	C: So I keep pushing myself and pushing myself and pushing myself but I'm getting now that it's getting too much to push myself. T: Ahh. What is it about that then? [Participant 8]

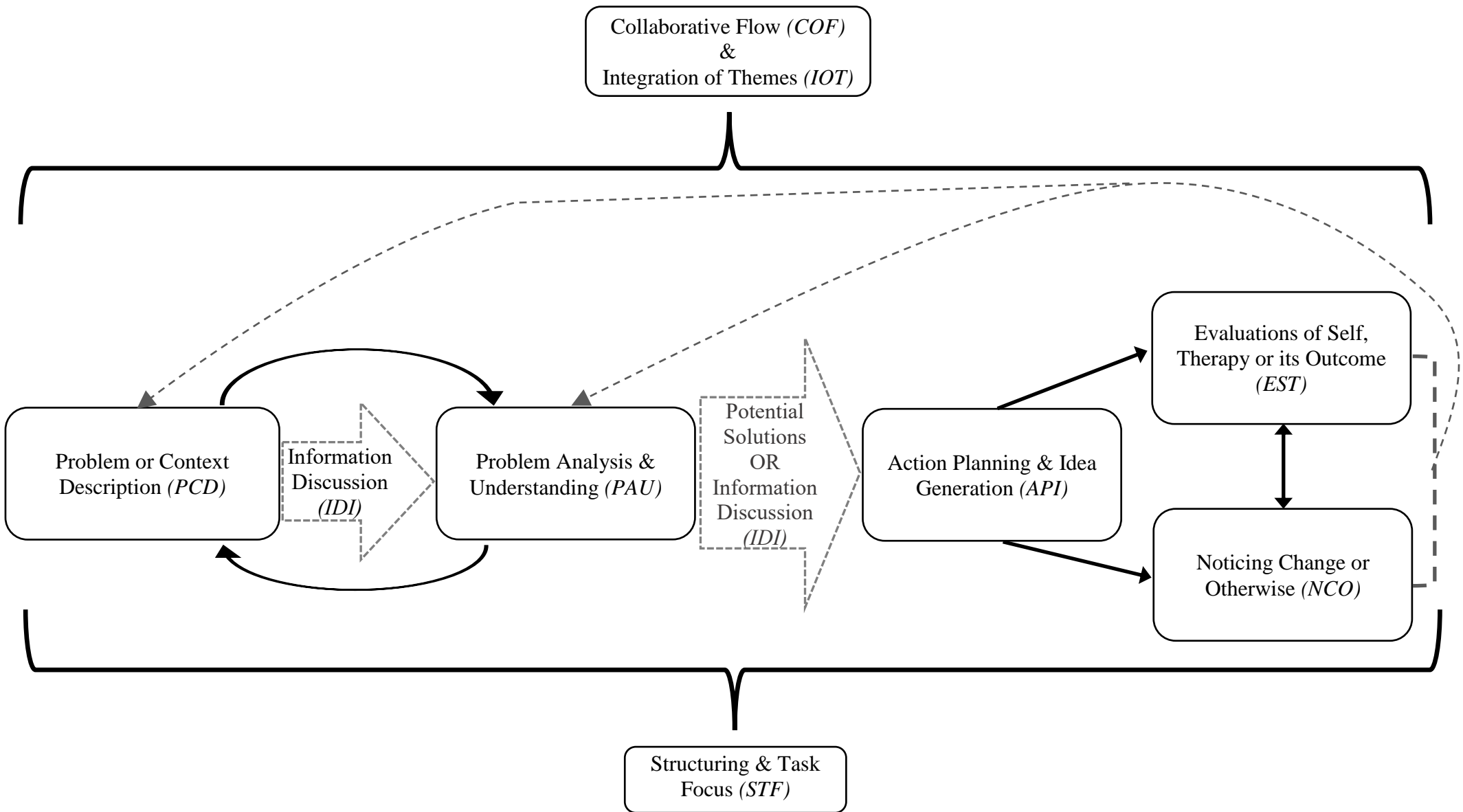|		
Problem analysis and understanding (PAU)	+2	**C:** I think as well because I think I am thinking differently, I'm not feeling as much pain. **T:** Really, right. **C:** I think I would have brought more on – because I get myself all tensed up and everything. [Participant 7]
	+1	**T:** Evaluating how worthwhile that the compensatory strategy is [ ] the fear isn't realised when you don't do it and you feel more confident and [ ] satisfied with yourself. **C:** I think I've always tried very, very hard to [ ] somehow change people's perception of me. Whereas I think I've come to realise that maybe the people that I was aspiring for to like me and to have those friendships with aren't the people that are probably naturally should be running after. [Participant 2]
	0	**T:** Where you've stayed with it long enough, that the anxiety starts to drop. That's, that's how you know. **C:** Okay, okay. [Participant 3]
	-1	**T:** I: So with those few factors in mind, your son staying at home during this week [ ], you're having your breakfast later and a different breakfast earlier [ ] and then I guess us meeting and talking about pain and testing it out. Are there any of those things…what of any of those things might have…? **C:** And I think also I mentioned to you last week my family have wanted me to have a wheelchair. [Participant 6]4]
	-2	**T:** While you have been drinking more though, you've had one night of full sleep. So what does that say? **C:** Yeah, mm, you're trying to say your system works and it doesn't. [Participant 9]
Structuring and task focus (STF)	+2	**C:** I also want to look at the goal setting thing as well, because I think that might be useful for me. **T:** OK. [Participant 5]
	+1	**T:** I was thinking it [ ] might be useful to have some kind of goal, general goals about (C: Yeah, definitely) what we're aiming for and that would be ok [ ]. **C:** Ok. [ ] **T:** We talked a bit about mindfulness [to] do that as well [ ] about half and half in terms [ ] of time. Is there anything [ ] other than that, that would you like to talk about? **C:** No that sounds pretty much probably what I was kind of expecting so that's cool. [Participant 2]
	0	**T:** OK, well what we were going to do today if that's OK as part of our session is kind of walk through the process that you have for dealing with your thrush, what you do to try to make it better and then when do you decide to go to the GP? So is that OK? **C:** Uh-huh. [Participant 1]
	-1	**T:** What did you go and see? **C:** Oh, the Woman in Black. **T:** Oh, that's quite scary. **C:** Really scary, yes, but I do like films that keep me on my toes. [Participant 3]
	-2	**C:** I give the neighbours a smoke signal [ ] I've got an incinerator – I give them a bit of smoke. I wait till they get their washing in – yeah. **T:** Oh well that's nice at least you will wait for the washing. **C:** Well yeah, you know the legal on that: there is no set time when you can have a bonfire as long as it isn't causing any danger. [Participant 9]

Collaborative Flow *(COF)*
&
Integration of Themes *(IOT)*

Problem or Context Description *(PCD)*

Information Discussion *(IDI)*

Problem Analysis & Understanding *(PAU)*

Potential Solutions OR Information Discussion *(IDI)*

Action Planning & Idea Generation *(API)*

Evaluations of Self, Therapy or its Outcome *(EST)*

Noticing Change or Otherwise *(NCO)*

Structuring & Task Focus *(STF)*

Table 1.

Participant primary care consultation rates and health problems.

ID	Face-to-face consultations in the previous 2 years[a]	Total CBT sessions	Total current morbidities[b]	Morbidity details
Participant 1	46	5	4	dyspepsia, IBS, atrial fibrillation, peripheral vascular disease
Participant 2	27	15	5	depression, GAD, IBS, psoriasis, eating disorder
Participant 3	39	40	5	chronic pain, asthma, CHD, GAD, IBS
Participant 4	28	22	3	depression, GAD, IBS
Participant 5	33	18	3	chronic pain, tendinitis, GAD
Participant 6	35	9	5	hypertension, myofascial pain syndrome, thyroid problems, IBS, glaucoma
Participant 7	28	6	3	hypertension, chronic pain, GAD
Participant 8	26	4	6	hypertension, arthritis, asthma, diabetes, CKD, IBD
Participant 9	60	4	6	hypertension, arthritis, diabetes, IBS, prostate disorder, IBD
Participant 10	35	5	5	hypertension, asthma, diabetes, arthritis, COPD,
Participant 11	29	11	8	hypertension, depression, diabetes, GAD, IBS, stroke, CKD, blindness
MEAN (SD)	*35 (10)*	*13(11)*	*5 (2)*	

Note. [a]The average UK primary care patient had 5.5 consultations per year at the time of data collection (Hippisley-Cox and Vinogradova, 2009).[b]Morbidities calculated through examination of medical records. Abbreviations: CHD, coronary heart disease; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary diseases; GAD, Generalized Anxiety Disorder; IBD, inflammatory bowel disease; IBS, irritable bowel syndrome.

Table 2.

Correlations between therapeutic alliance, therapist competence and CICS themes.

Consultation Interaction Coding Scheme theme	WAI-O Correlation *p value*	CTS-R Correlation *p value*
Action Planning and Idea Generation (*API*)	.72** *<.001*	.39 *.071*
Evaluations of Self or Therapy (*EST*)	.83** *<.001*	.73** *<.001*
Information Discussion (*IDI*)	.08 *.719*	-.03 *.883*
Noticing Change or Otherwise (*NCO*)	.83** *<.001*	.40 *.066*
Problem Analysis and Understanding (*PAU*)	.80** *<.001*	.55** *.008*
Problem and Context Description (*PCD)*	-.71** *<.001*	-.47* *.029*
Structuring and Task Focus (*STF*)	.36 *.102*	.41 *.058*
Collaborative Flow (*COF*)	.89** *<.001*	.57** *.006*
Integration of Themes (*IOT*)	.84** *<.001*	.68** *<.001*
Global average	.90** *<.001*	.63** *.002*
Content average	.90** *<.001*	.62** *.002*
WAI-O	- *-*	.67** *.001*

Note. Abbreviations: WAI-O, working alliance inventory – observer-rated; CTS-R, cognitive therapy rating scale – revised. ** p < .01; * p < .05 level (2-tailed).