

A Coding System to Measure Elements of Shared Decision Making During Psychiatric Visits

Michelle P. Salyers, Ph.D.  
1481 W. 10<sup>th</sup> Street  
Indianapolis, IN 46202  
[mpsalyer@iupui.edu](mailto:mpsalyer@iupui.edu)  
(317)988-4419  
Fax (317)988-2719

Marianne S. Matthias, Ph.D.  
1481 W. 10<sup>th</sup> Street  
Indianapolis, IN 46202  
[mmatthia@iupui.edu](mailto:mmatthia@iupui.edu)  
(317)988-4514  
Fax (317)988-2719

Sadaaki Fukui, Ph.D.  
The University of Kansas Center for Research Methods and Data Analysis & School of Social  
Welfare Office of Mental Health Research and Training  
1545 Lilac Lane  
Lawrence, KS 66044  
[fsadaaki@ku.edu](mailto:fsadaaki@ku.edu)  
(785)864-5874  
Fax (785)864-5277

Mark C. Holter, Ph.D.  
University of Kansas  
School of Social Welfare  
Twente Hall  
Office of Mental Health Research and Training  
1545 Lilac Lane  
Lawrence, KS 66044  
[mholter@ku.edu](mailto:mholter@ku.edu)  
(785)864-4720  
Fax (785)864-5277

Linda Collins, B.S.  
1481 W. 10<sup>th</sup> Street  
Indianapolis, IN 46202  
[Linda.collins4@va.gov](mailto:Linda.collins4@va.gov)  
(317)988-2722  
Fax (317)988-2719

Nichole Rose, M.S.  
ACT Center of Indiana  
402 North Blackford, LD 110  
Indianapolis, IN 46202

[nirose@uemail.iu.edu](mailto:nirose@uemail.iu.edu)  
(812)371-4986

John Thompson, M.A.T.S., M.S.W.  
University of Kansas  
School of Social Welfare  
Twente Hall  
Office of Mental Health Research and Training  
1545 Lilac Lane  
Lawrence, KS 66044  
[johnyt@ku.edu](mailto:johnyt@ku.edu)  
(785)864-4720  
Fax (785)864-5277

Melinda Coffman, B.S.  
University of Kansas  
School of Social Welfare  
Twente Hall – Rm. 3  
Office of Mental Health Research and Training  
1545 Lilac Lane  
Lawrence, KS 66044  
[coffmanm@ku.edu](mailto:coffmanm@ku.edu)  
(785)864-5868  
Fax (785)864-5277

William C. Torrey, M.D.  
Dartmouth-Hitchcock Medical Center  
Department of Psychiatry  
One Medical Center Drive  
Lebanon, NH 03756  
[William.C.Torrey@Dartmouth.EDU](mailto:William.C.Torrey@Dartmouth.EDU)  
(603)-650-6069  
Fax (603)-650-5842

Running Head: Measuring Shared Decision Making

Accepted for publication in: *Psychiatric Services*

**Role of Funding Source**

Funding for this study was provided by NIMH Grant (R24 MH074670; Recovery Oriented Assertive Community Treatment), by the VA HSR&D Center of Excellence in Implementing Evidence-based Practices, and by the Kansas Department of Social and Rehabilitation Services. The funders had no further role in study design, in the collection, analysis and interpretation of data, in the writing of the report, and in the decision to submit the paper for publication.

### **Acknowledgements**

We thank providers and consumers from Adult & Child Center (Indianapolis, IN), Wyandot (Kansas City, KS), and Bert Nash (Lawrence, KS). We also thank Candice Hudson and Sylwia Oles for their assistance.

Accepted version; Final version published as Salyers, M. P., Matthias, M. S., Fukui, S., Holter, M. C., Collins, L., Rose, N., Thompson, J., Coffman, M., Torrey, W. C. (2012). A coding system to measure elements of shared decision making during psychiatric visits, *Psychiatric Services*, 63, 779-784.

Measuring Shared Decision Making in Psychiatric Care

Abstract

*Objective:* Shared decision making is widely recognized to facilitate effective health care; tools are needed to measure the level of shared decision making in psychiatric practice. *Methods:* A coding scheme assessing shared decision making in medical settings (1) was adapted, including creation of a manual. Trained raters analyzed 170 audio recordings of psychiatric medication check-up visits. *Results:* Inter-rater reliability among three raters for a subset of 20 recordings ranged from 67% to 100% agreement for the presence of each of nine elements of shared decision making and 100% for the overall agreement between provider and consumer. Just over half of the decisions met minimum criteria for shared decision making. Shared decision making was not related to length of visit after controlling for complexity of decision. *Conclusions:* The shared decision making rating scale appears to reliably assess shared decision making in psychiatric practice and could be helpful for future research, training, and implementation efforts.

Shared decision making is a collaborative process between a provider and a consumer of health services that entails sharing information and perspectives, and coming to an agreement on a treatment plan (2, 3). This collaborative process is viewed as central to high-quality, patient-centered healthcare and has been identified as one of the top ten elements to guide the redesign of healthcare (4). Shared decision making is also a growing area of interest in psychiatry, particularly in the treatment of severe mental illnesses (5-8). However, in order to study shared decision making and ensure its widespread use, tools are needed to assess whether core elements are present.

We have adapted a widely used tool for evaluating shared decision making, based on the work of Braddock and colleagues (1, 9-13), for use in psychiatry. This coding scheme is comprehensive and takes into account both providers' and consumers' input. Although we considered the OPTION scale, another reliable approach to coding observed medical visits (14, 15), scoring is based solely on the provider behavior. Given the important role consumers should play in the decision making process (3, 16), an ideal scoring system should also account for active consumer involvement.

The purpose of this study was to assess the applicability of Braddock's coding system to psychiatric visits. We tested inter-rater reliability and examined the frequency with which elements of shared decision making were observed. Given concerns about time constraints as potential barriers for shared decision making in psychiatry (8) and the general medical field (17), we also explored whether consultations with high levels of shared decision making would take more time.

## Methods

### *Study Design and Sample*

We combined audio recorded psychiatric visits from three prior studies that took place between September 2007 and April 2009. Participants were prescribers and adult consumers of community mental health centers in either [state] or [state]. Study 1 was an observational study of 40 psychiatric visits (four providers, with ten consumers each) examining how consumers with severe mental illness may be active in treatment sessions (16). Study 2 was a randomized intervention study, examining the impact of the Decision Support Center (16) to improve shared decision making in medication consultations. We used baseline recordings, prior to intervention (three providers and 98 consumers). Study 3 was an observational study of psychiatric visits with one provider and 48 consumers. In all, the sample included eight providers (five psychiatrists and three nurse practitioners) and 186 consumers. Because of recording difficulties, 170 consumers had usable audio recordings.

The sample was predominantly Caucasian (54%), with 40% African-American, and 6% reporting another race. Approximately half were male (52%), and the mean age was  $43.6 \pm 11.2$  years. Diagnoses included schizophrenia spectrum disorders (55%), bipolar disorder (22%), major depression (14%), and other disorders (9%).

### *Measures*

Background Characteristics. In study 1, participants reported demographics in a survey prior to the recorded visit; providers reported mental health diagnoses. In Studies 2 and 3, demographics and diagnoses were obtained from a statewide automated information database.

Shared Decision Making. We adapted the Elements of Informed Decision Making Scale (10). Based on recordings and/or transcripts from medical visits, trained raters identify whether a clinical decision is present (i.e., a verbal commitment to a course of action addressing a clinical issue), and classify the type of decision as basic, intermediate, or complex. Basic decisions are expected to have minimal impact on the consumer, high consensus in the medical community, clear probable outcomes, and pose little risk (e.g., deciding what time to take a

medication). Intermediate decisions have moderate impact on the consumer, wide medical consensus, moderate uncertainty, and may pose some risk to the consumer (e.g., prescribing an antidepressant). Complex decisions may have extensive impact on the consumer, have uncertain outcomes and/or controversy in the medical literature, and pose a risk to consumers (e.g., prescribing Clozapine). For each decision, raters determine the presence of nine elements: consumer's role in decision making, consumer's context (i.e., how the problem/decision may impact their life), the clinical nature of the decision, alternatives, pros and cons, uncertainties and/or likelihood of success, assessment of consumer understanding, consumer's desire to involve others in the decision, and assessment of consumer preferences. Each element is rated as absent = 0, partial = 1 (brief mention of the topic), or complete = 2 (reciprocal discussion, with both parties commenting). Items are summed for an overall score ranging from 0 to 18.

Braddock et al. (1) also describe a minimum level of decision making based on the presence of specific elements according to the complexity of a decision. For basic decisions, minimum criteria for shared decision making include the clinical nature of the decision (element #3) and either the consumer's desired role in decision making (#1) or the consumer's preference (#9). For intermediate decisions, minimum criteria for shared decision making additionally requires alternatives (#4), discussing pros and cons of the alternatives (#5), and assessing the consumer's understanding (#7). Complex decisions require all nine elements. Braddock's coding system has been used in several studies of decision making with primary care physicians and surgeons with high reliability (1, 9, 10, 12). We use the term SDM-18 for the sum of items and SDM-Min for the minimum criteria for shared decision making.

Adapting the measure for psychiatry. Initially we developed a codebook from published descriptions (1, 9, 10) and later reviewed Braddock's codebook for further clarification. An initial team (psychologist, psychiatrist, health communication expert, and two research assistants)

read individual transcripts, applied codes, and met to develop examples from our transcripts. This was an iterative process of individual coding followed by consensus discussions. We kept all nine elements, but added an additional code to “alternatives” (element #4) to classify whether non-medication alternatives were discussed. We believed this was important as medications can interfere with other activities consumers do to maintain wellness (18). We also added ratings to assess who initiated each element (to better identify consumer activity) and an overall rating to classify the level of agreement about the decision between provider and consumer (e.g., full agreement of both parties, passive/reluctant agreement by consumer or provider, disagreement by consumer or provider). These aspects are not part of the overall shared decision making score, but provide descriptive information. Once we achieved reliable agreement, we trained two additional raters using the codebook and the initial transcripts. Throughout this process, we modified the manual, adding clarifications and additional examples as appropriate.

### *Coding Procedures*

All transcripts were given a random ID number and identifying information was removed, including references to location so that raters would be blind to setting. Raters coded transcripts individually. Every two weeks, we distributed a transcript to be coded by all raters, followed by a consensus discussion, to maintain consistent coding.

### *Data Analyses*

We evaluated inter-rater agreement among three coders (a psychologist and two research assistants) for 20 randomly selected transcripts. Agreement was assessed by both percentage agreement and Gwet’s agreement coefficient (AC1; (19)). Although Kappa is often used, it does not correct for chance agreement, and is difficult to use with multiple raters (19). Gwet’s AC1 allows for the extension to multiple raters and multiple category responses, and adjusts for chance agreement and misclassification errors. Given other work with AC1 (20),



coefficients above .8 can be considered very strong, .6-.8 moderate, and .3-.5 fair agreement. We also present percent agreement, because all possible response categories are not necessarily observed in the 20 recordings, which can leave corresponding cell frequencies empty. Hence, interpretations of the inter-rater agreement involve both percent agreement and AC1. We present descriptive data on SDM-18, SDM-Min, who initiated each element, and the overall agreement between the provider and consumer in the decision. Finally, we examined the relationship between shared decision making scores and length of time in the session, controlling for level of decision complexity. These analyses involved partial correlation for SDM-18 scores and analysis of co-variance for SDM-Min. All procedures for this study were approved by the Institutional Review Board at [university].

## Results

Inter-rater reliability across the elements of shared decision making was strong (Table 1). Percent agreement ranged from 67% (discussion of the consumer's role in decision making) to 100% (discussion of consumer's context). The AC1 statistic was moderate to excellent for all elements (AC1=.68-.97) except consumer's role in decision making (AC1 = .51). Agreement for who initiated each element ranged from 73% (context) to 100% (clinical nature); AC1 statistics were moderate to excellent, ranging from .66 to .97. There was 100% agreement among raters on the classification of the final agreement between consumer and provider regarding course of action for the decision.

Overall, 128 of the 170 sessions (75%) contained a clinical decision. Types included: stopped a medication (6%), added a medication (14%), changed the time or administration of a previously prescribed medication (14%), changed the dosage of a medication (21%), decided not to change a medication when an alternative was offered (30%), or decided on a non-medicinal alternative (37%). More than one of these decisions may have been present in the

same discussion. However, we coded the shared decision making elements on the basis of the overall discussion because of the highly related nature of the decisions (e.g., decreasing one medication and adding a new medication to address a symptom). Only one visit contained two clearly separate clinical issues resulting in decisions, which were scored separately with the first decision included in these analyses. Overall, our sample included primarily basic [n=59, 46%] or intermediate [n=67, 52%] decisions, with only two decisions rated as complex [2%].

The frequencies of observed elements of shared decision making and who initiated them are shown in Table 2. Decisions most often included a complete discussion of the consumer's context, i.e., how his or her life is being impacted by the clinical concern (92%). Decisions also frequently included discussions of the clinical nature of the decision (63%), alternatives to address the concern (58%; notably over half of these also included non-medication alternatives), and the consumer's preference (56%). The elements that occurred least frequently were assessment of consumer's desire for others' input (6%) and assessment of consumer's understanding (7%). In terms of who initiated discussion of the elements, the provider was the primary initiator of all elements but one. Consumers most often initiated a discussion about the context of the clinical concern (66% of the time).

Agreement between provider and consumer was high -- 101 decisions (79%) were rated as being in full agreement. We observed the consumer passively or reluctantly agreeing in 15% of decisions and the provider passively or reluctantly agreeing in 6% of decisions. We observed no decisions in which the consumer or provider disagreed with the final decision.

Finally, we examined length of visit in relation to shared decision making. Mean length of time and shared decision making scores are shown by type of decision in Table 3. Across all decisions, the mean length of visits was  $16.8 \pm 7.0$  minutes (range: 3-36) and the mean SDM-18 was  $9.7 \pm 3.3$  (range: 2-17). The bivariate correlation between them was  $r = .25$ ,  $df = 124$ ,  $p < .01$ .

However, after controlling for complexity of the decision, the partial correlation between length of time in the visit and SDM was no longer significant. In addition, analysis of co-variance revealed no significant effect of SDM-Min on length of time, controlling for level of complexity of the decision.

### Discussion

The shared decision making rating scale appears to be a reliable tool for assessing the level of shared decision making in psychiatric visits. Our raters achieved moderate to strong levels of agreement on individual elements of shared decision making, who initiated them, and the overall agreement between provider and consumer. In addition, the codebook created from this work, with examples from actual psychiatric visits, could be a useful tool for others seeking to measure shared decision making in psychiatric practice.

In terms of individual items, the element with the lowest and only “fair” reliability was the consumer’s role in decision making. This may be a function of our scoring procedures. While scoring, we were aware that either this element or the consumer’s preference could count toward the overall decision making score. In our consensus meetings we often discussed whether particular quotes from a transcript should be included as evidence for a role in decision making or for preference about the decision. Because either would count towards an overall decision making score, we were less concerned about lack of agreement on the role item. To increase ease of use of the scale and to enhance reliability, we recommend integrating these items, for example, by including “role” as partial credit for an overall “preference” item.

We found that consumers and providers fully agreed on the decision 79% of the time, at least as judged by raters on the basis of statements in the transcripts. We have no way to determine how consumers or providers perceived the decisions. In addition, although overall agreement on the course of treatment appeared strong, just over half of decisions (53%) in our

sample met minimum criteria for shared decision making. When examined by decision complexity, 61% of the basic decisions met minimum criteria, compared to only 46% of the intermediate or complex decisions. These rates were similar to the orthopedic surgery sample (10) and higher than the mixed sample of primary care and surgery (1), though they are certainly less than ideal.

Notably, our sample of decisions (as well as the prior samples) rarely contained an assessment of the consumer's understanding or discussed the consumer's desired role in decision making and/or desired role of others in helping with the decision. These appear to be important areas for growth in order to ensure a truly shared and fully informed decision making process. In addition, two other areas were frequently absent (30% of the time or more): discussion of uncertainties regarding the decision and discussion of the pros and cons of the decision. The absence of these elements is particularly concerning in the context of intermediate and complex decisions. Tools, like written or electronic decision aids can enhance consumer involvement and knowledge regarding treatment decisions (21). In psychiatric settings, electronic decision tools appear particularly promising (22, 23).

We found a high level of consumer context discussed -- 92% were "complete," with reciprocal sharing of information. Psychiatry may generally include psychosocial aspects of clinical problems and treatment. On one hand, a high prevalence is a positive finding because, at least in our sample, one element of shared decision making is being incorporated almost all of the time. On the other hand, given the high level of scoring, this item may not be useful to meaningfully distinguish among decisions. Perhaps by including more specific attention to consumer goals (rather than just broader life context), this item could be more sensitive to important variations in practice. This shift to explicit attention to consumer goals would be consistent with recovery-oriented principles of care (18, 24). In addition, it may be useful to code

both for broader context as well as for specific consumer goals (e.g., consumer wants to work and is concerned about sedating side effects).

In our adaptation of the shared decision making coding system, we added a rating to describe who initiated each element as a potential way to identify how active consumers were in the decision making process. In this initial sample, consumers appeared to be the primary initiator of discussions regarding life context, whereas providers were primary initiators of all other elements. The elements where consumers appeared somewhat active (e.g., initiating discussion over 30% of the time) included discussing non-medication alternatives, checking their own understanding, and stating preferences. Had we rated the clinician's behavior alone, e.g., with the OPTION scale (14), we may have missed some of these aspects of shared decision making. In addition, these elements may be fruitful areas to focus efforts on increasing consumer partnership; for example, interventions to coach people with mental illness to ask more questions (25).

Shared decision making was correlated with longer visit time. Similarly, Goss and colleagues found high shared decision making using the OPTION scale correlated with longer visits (26). However, they did not control for complexity of decision. In our sample, intermediate and complex decisions took more time than basic decisions. After accounting for this, shared decision making was not related to length of time of the visit, a finding consistent with Braddock and colleagues (1, 10). Given that physicians frequently cite increased time as an obstacle to shared decision making (17), these findings provide additional rationale for the feasibility of shared decision making in psychiatric settings.

We adapted Braddock's scale to psychiatric visits by using a convenience sample of recorded psychiatric visits in community mental health settings. The sample was small, and included providers willing to be audio recorded, some of whom were beginning participation in a

study to enhance shared decision making. The rates of shared decision making seen in this sample may not generalize to other consumers and providers; it is notable, however, that even in this willing sample, rates of shared decision making were still modest. Further work is needed in more diverse samples to fully establish generalizability of the rating system. In addition, although reliability was strong and the scale appears applicable to psychiatric settings, we were not able to test criterion-related validity of the scale, for example, predicting consumer satisfaction, treatment concordance, or other possible outcomes of a more shared process in decision making. Future studies could also examine predictors and possible moderators of shared decision making in longitudinal designs. For example, length of time seeing the same provider and how long the particular problem had been addressed could alter the frequency of the shared decision making elements observed. Despite these limitations, this scale is a promising approach to measuring shared decision making in psychiatric visits and an important step in advancing the study and application of shared decision making in mental health care.