

Online Supplement DS1

The Field Trial Version of the Core Cultural Formulation Interview

1. What problems or concerns bring you to the clinic?
2. What troubles you most about your problem?
3. People often understand their problems in their own way, which may be similar or different from how doctors explain the problem. How would you describe your problem to someone else?
 - 3a. Sometimes people use particular words or phrases to talk about their problems. Is there a specific term or expression that describes your problem?
 - 3b. What is it?
4. Why do you think this is happening to you? What do you think are the particular causes of your problem?
5. What, if anything, makes your problem worse, or makes it harder to cope with?
 - 5a. What have your family, friends, and other people in your life done that may have made your problem worse?
6. What, if anything, makes your problem better, or helps you cope with it more easily?
 - 6a. What have your family, friends, and other people in your life done that may have made your problem better?
7. Is there anything about your background, for example your culture, race, ethnicity, religion or geographical origin that is causing problems for you in your current life situation? In what way?
8. On the other hand, is there anything about your background that helps you to cope with your current life situation? In what way?
9. Sometimes people consider various ways of making themselves feel better. What have you done on your own to cope with your problem?
10. Often, people also look for help from other individuals, groups, or institutions to help them feel better. In the past, what kind of treatment or help from other sources have you sought for your problem?
 - 10a. What type of help or treatment was most useful? Why?/How?
 - 10b. What type of help or treatment was not useful? Why?/How?
11. Has anything prevented you from getting the help you need—for example, cost or lack of insurance coverage, getting time off work or family responsibilities, concern about stigma or discrimination, or lack of services that understand your language or culture? What got in the way?
12. Now let's talk about the help you would be getting here. Is there anything about my own background that might make it difficult for me to understand or help you with your problem?
 - 12a. In what way?/Why not?
13. How can I and others at our clinic be most helpful for you?
14. What kind of help would you like from us now, as specialists in mental health?

Online Supplement DS2

Reliability of the Debriefing Instrument for Patients and the Debriefing Instrument for Clinicians

The Debriefing Instrument for Patients (DIP) and the Debriefing Instrument for Clinicians (DIC) are each composed of three domains that assess respondents' perceptions of the feasibility, acceptability, and clinical utility of the Cultural Formulation Interview. We estimated the reliability (both raw and standardized coefficient α) of each domain prior to calculating mean DIP and DIC scores. While the DIC items appeared reasonably reliable as written, the DIP items showed greater variation; psychometric evaluation led to the exclusion of two items, one in the feasibility domain and one in the acceptability domain. This document describes the procedures we followed to reach this conclusion. The psychometric analyses of the DIP clinical utility domain and all three DIC domains are not presented in this document, since no differences between raw and standardized α s were observed; all of the domains had adequate reliability, item correlations with total, inter-item correlations, and changes to α by item; and no items in these domains were changed or removed. The final DIP and DIC item-based results and α coefficients are presented in Table DS1 below.

Table DS1. Final domain items, means, and reliability estimates for Debriefing Instrument – Patients and Clinicians¹

Patient interviews			Clinician interviews		
Domains and items	Mean	SD	Domains and items	Mean	SD
Feasibility (n = 302) $\alpha = 0.45$ (Raw = 0.45)			Feasibility (n = 312) $\alpha = 0.78$ (Raw = 0.77)		
09. Were easy to understand	1.37	0.72	12. Were easy to administer	0.97	0.97
11. Improved the flow of the interview	1.30	0.71	13. Were easily understood by the patient	0.56	1.14
			14. Contributed positively to the flow of my clinical interview	0.77	1.11
Acceptability (n = 299) $\alpha = 0.48$ (Raw = 0.49)			Acceptability (n=297) $\alpha = 0.80$ (Raw=0.79)		
13. Should be asked by every clinician.	1.14	0.99	15. Helped make the patient feel more at ease during the interview	0.91	1.01
14. Helped me feel more at ease with the interview	1.40	0.75	16. Can be incorporated by mental health clinicians into routine clinical interviews	1.06	0.87
			17. Facilitated a good assessment of cultural factors relevant to clinical care	0.95	1.01
			18. I would recommend for use by other mental health clinicians	1.08	0.84
Clinical Utility (n = 275) $\alpha = 0.82$ (Raw = 0.82)			Clinical Utility (n = 290) $\alpha = 0.89$ (Raw=0.89)		
01. Helped me explain my main concerns	1.48	0.56	01. Helped me understand the patient's cultural background	0.74	1.11
02. Helped me communicate important aspects of my background, such as religious faith and/or culture	1.20	0.83	02. Clarified the patient's ideas about the cause of the problem	0.98	1.01
03. Helped me understand how	1.20	0.85	03. Clarified my understanding of the	0.95	1.03

my background and current situation affect my problem			patient's symptoms and problems		
04. Helped me explain what kinds of help I would like	1.36	0.69	04. Gave me confidence in the diagnosis	0.58	1.24
05. Gave me confidence that the clinician understood my situation	1.50	0.68	05. Facilitated treatment planning	0.98	1.06
06. Helped me identify things that could get in the way of my treatment	1.05	0.97	06. Helped me identify issues that could interfere with treatment adherence	1.11	0.95
07. Encouraged me to share important information that might not have been mentioned otherwise	1.21	0.93	07. Helped me identify additional aspects or dimensions of the patient's clinical problems	1.04	1.01
08. Were useful overall	1.44	0.62	08. Helped me assess the severity of the patient's clinical problems	0.77	1.11
			09. Facilitated my rapport with the patient	1.16	1.02
			10. Clarified how my perspective on the patient's presentation was similar or different to the patient's	0.79	1.10
			11. Were useful overall	1.12	0.75

¹. Standardized α 's are reported in bold, with raw α 's in parentheses

All reliability calculations were conducted with SAS Version 9.4 (Cary, NC) using the CORR procedure with the ALPHA and NOMISS options. Only patients and clinicians who answered every item within a domain were included for the reliability analyses. Sample sizes increased between calculations done with the full DIP domains and the reduced DIP domains because the former excluded patients who did not provide answers for items that were ultimately removed (DIP item 10 and 12). Alpha estimates and item means were recalculated after these items were identified and removed. Subsequent domain scores are based on these reduced scales.

DIP DOMAINS ON FEASIBILITY AND ACCEPTABILITY

The Feasibility and Acceptability domain each included one negatively-worded item intended to be scored in reverse (items 10 and 12; Table DS2); these two domains had the lowest standardized α s (0.18 and 0.17, respectively) after applying the reverse-scoring scheme. There was also a large difference between the raw and standardized α s of these two domains (0.07 and 0.18; 0.07 and 0.17, respectively), suggesting that the variance of at least one item within each domain was appreciably different from that of the other items (DeVellis, 2012).¹ The Clinical Utility domain produced acceptable α values with no difference between the raw and standardized versions (both α s=0.82) and is included fully in Table DS1. Such low α s for the Feasibility and Acceptability domains warranted a search for problematic items.

Table DS2. Debriefing Instrument for Patients (DIP), original domain composition using all items: reliability estimates and item means.

Patient interviews			
Domains and items	Domain α	Mean	SD
Feasibility (n=298)	Raw: 0.07 Standardized: 0.18		
09. Were easy to understand		1.36	0.72
10. Took more time to share my perspective than I wanted. ¹		-0.20	1.33
11. Improved the flow of the interview		1.30	0.72

Acceptability (n=295)	Raw: 0.07		
	Standardized: 0.17		
12. Were too personal. ^r		0.28	1.23
13. Should be asked by every clinician.		1.14	1.00
14. Helped me feel more at ease with the interview		1.41	0.74

^r. Reverse-scored

Feasibility: Table DS3A lists the correlation of each item of the Feasibility domain with the total correlation of the remaining items. DIP 10 clearly stands out as unusual, producing a weak negative correlation with the other items despite reverse-coding the negatively-worded item.

Table DS3. DIP individual item correlation with total, by domain

A.

DIP Feasibility	Correlation with Total	
	Raw	Standardized
DIP 09	0.197065	0.262147
DIP 10 ^r	-0.053345	-0.053513
DIP 11	0.013819	0.102627

^r. Reverse-scored

B.

DIP Acceptability	Correlation with Total	
	Raw	Standardized
DIP 12 ^r	-0.088108	-0.075912
DIP 13	0.044750	0.130489
DIP 14	0.215033	0.246186

^r. Reverse-scored

Table DS4A examines the inter-item correlations more closely, displaying the correlation matrix among the three items in the domain. Item 09 appears to be moderately correlated with item 11, but weakly correlated with item 10. Item 11 also appears to be poorly correlated with item 10, as well as negatively correlated.

Table DS4. DIP inter-item correlation matrix, by domain

A.

DIP Feasibility	DIP 09	DIP 10 ^r	DIP 11
DIP 09	1.000	0.05439	0.28935
DIP 10 ^r		1.000	-0.1403
DIP 11			1.000

^r. Reverse-scored

B.			
DIP Acceptability	DIP 12 ^r .	DIP 13	DIP 14
DIP 12 ^r .	1.000	-0.13121	-0.00803
DIP 13		1.000	0.31649
DIP 14			1.000

Table DS5 presents changes to α values if an item were dropped. The α of the Feasibility domain (Table DS5A) increases substantially when item 10 is removed (raw and standardized $\alpha=0.44$), providing further evidence that item 10 may not fit with the other two items. If any other item were to be removed, the α value decreases further (dropping item 11) or results in a negative value (item 09). In the latter case, the hypothetical domain composed of items 10 and 11 produces a negative α , suggesting that the items may not be measuring the same construct. The fact that the raw and standardized α s are essentially identical after dropping item 10 also suggests that the variance between item 09 and 11 is similar.

Table DS5. Original DIP coefficient α after deleted item

A.		
DIP Feasibility		
Deleted Item	Raw α	Standardized α
DIP 09	-0.27	-0.33
DIP 10 ^r .	0.45	0.45
DIP 11	0.09	0.10

B.		
DIP Acceptability		
Deleted Item	Raw α	Standardized α
DIP 12 ^r .	0.46	0.48
DIP 13	0.01	0.02
DIP 14	-0.29	-0.30

For these reasons we excluded item 10 from DIP-Feasibility scoring. Including the patients who were excluded from the full DIP-Feasibility reliability calculation due to missing item 10, the final standardized α estimate was 0.45.

Acceptability: In the Acceptability domain the reverse-scored item was also problematic. Item 12 correlated poorly and negatively with the total remaining items (Tables DS3B and DS4B). Alpha also improved substantially once the item was removed (Table DS5B). Removing any other item resulted in an α nearly at 0 (dropping Item 13) or a negative α (dropping Item 14). As in the Feasibility domain, it appears that the reverse-scored item may not be measuring the intended construct. Dropping this item also produced very similar raw and standardized α estimates.

Therefore, we excluded item 12 from DIP-Acceptability scoring. Including the patients who were excluded from the original reliability calculation for missing item 12, the final standardized α for DIP-Acceptability was 0.48.

Additional reference

1. DeVellis RF. Scale development: Theories and applications. 3rd ed. Thousand Oaks, CA: Sage Publications, Inc.; 2012.

Related Characteristics^f

Foreign birth	32	96.97	17	56.67
State of birth				
Andhra Pradesh	1	0.99		
Assam	1	0.99		
Bihar	11	10.89		
Gujarat	2	1.98		
Haryana	3	2.97		
Himachal Pradesh	1	0.99		
Madhya Pradesh	3	2.97		
Maharashtra	32	31.68		
National Capital/ Territory of Delhi	29	28.71		
Not born in India	1	0.99		
Punjab	1	0.99		
Rajasthan	5	4.95		
Tamil Nadu	1	0.99		
Uttar Pradesh	8	7.92		
Uttarakhand	2	1.98		
Tribe				
Arab		2	6.90	
Kalenjin		1	3.45	
Kamba		5	17.24	
Kikuyu		13	44.83	
Kisii		1	3.45	
Luhya		3	10.34	
Luo		1	3.45	
Somali		2	6.90	
Taita		1	3.45	
Race				
Mixed, primarily indigenous			14	41.18
Mixed, primarily black			2	5.88
Mixed, primarily white			16	47.06

Mixed, primarily Asian	1	2.94												
White non-Hispanic	1	2.94												
Race/Ethnicity														
Hispanic											54	60.00 ^a		
Non-Hispanic white											13	14.44 ^a		
Non-Hispanic black											5	5.56 ^a		
Non-Hispanic American Indian											1	1.11 ^a		
Non-Hispanic East Asian											14	15.56 ^a		
Non-Hispanic South Asian											2	2.20 ^a		
Mixed/Other											1	1.11 ^a		

Number of Patients with at Least One Diagnosis in Disorder Cluster^g

Anxiety Disorders	6	18.18	21	20.79	0	0	17	56.67	11	32.35	28	30.77	83	26.10
Bipolar Disorders	1	3.03	8	7.92	7	24.14	0	0	2	5.88	14	15.38	32	10.06
Depressive Disorders	23	69.70	33	32.67	3	10.34	22	73.33	19	55.88	46	50.55	146	45.91
Psychotic Disorders	1	3.03	11	10.89	15	51.72	3	10.00	2	5.88	25	27.47	57	17.92
Substance Disorders	3	9.09	9	8.91	4	13.79	2	6.67	4	11.76	14	15.38	36	11.32
Other Disorders	3	9.09	20	19.80	3	10.34	6	20.00	9	26.47	10	10.99	51	16.04

a. Data unavailable for 1 participant.

b. Fular, Kirundi, Kiswahili, Moroccan, Moroccan Arabic, Rwandese, and Wolof

c. Cantonese and Mandarin

d. Gujarati, Hindi, Marathi, Punjabi, Tamil, Telgu, and Urdu

e. Arabic, Armenian, Bosnian, Dari, French, Hmong, Indonesian, Ingushetian, Kurdish, and Turkish

f. As there is no standard for reporting race and ethnicity in international trials, we instead report salient demographic factors as identified by local sites and recognized by governments. Therefore, not all factors will be relevant to every country.

g. Diagnoses made after conducting CFI and a diagnostic interview. Patients can have multiple diagnoses so percentages will sum to over 100%.

*p<0.05; **p<0.01, ***p<0.001.

Table DS7 Cross-national comparison of feasibility, acceptability, and clinical utility of the CFI

	Canada (n=33)	India (n=101)	Kenya (n=30)	Netherlands (n=30)	Peru (34)	USA (n=91)	Overall (n=315)	Statistic F(5)	p-value
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)		
Patients									
Feasibility	1.50 (0.51)	1.32 (0.47) ^a	1.75 (0.42) ^{a,b,c,d}	1.02 (0.66) ^{c,e}	1.21 (0.70) ^d	1.30 (0.58) ^{b,e}	1.33 (0.57)	4.32	0.001**
Acceptability	1.23 (0.77) ^a	1.21 (0.67) ^b	2.00 (0) ^{a,b,c,d,e}	0.76 (0.83) ^{c,f}	1.23 (0.67) ^d	1.295 (0.63) ^{e,f}	1.27 (0.71)	9.34	<0.001***
Clinical Utility	1.49 (0.43)	1.17 (0.49)	1.56 (0.34) ^a	0.99 (0.56) ^a	1.29 (0.42)	1.27 (0.60)	1.26 (0.53)	2.60	0.026*
Clinicians									
Feasibility	0.24 (0.83)	0.88 (0.91)	1.36 (0.49)	0.52 (0.94)	0.62 (0.76)	0.75 (0.92)	0.75 (0.90)	2.30	0.046*
Acceptability	0.44 (0.85)	0.96 (0.80)	1.47 (0.50)	0.66 (0.69)	1.07 (0.52)	1.09 (0.65)	0.98 (0.75)	2.13	0.063
Clinical Utility	0.27 (0.81) ^{a,b,c}	1.02 (0.65) ^a	1.45 (0.43) ^b	0.58 (0.59)	1.28 (0.43) ^c	0.90 (0.67)	0.93 (0.70)	4.26	0.001**
Duration (min)									
CFI	26.89 (7.73)	23.05 (10.89)	29.21 (3.05) ^{a,b}	18.82 (8.78) ^a	19.68 (8.59) ^b	23.69 (9.24)	23.41 (9.57)	3.12	0.01*
Total Interview	84.39 (26.45) ^{a,b,c,d}	43.43 (17.07) ^{b,e}	37.57 (3.80) ^{c,f}	88.18 (29.20) ^{e,f,g,h}	37.83 (13.52) ^{d,g}	54.04 (16.68) ^{a,h}	54.12 (25.61)	22.46	<0.001***
CFI proportion of total interview	35.20% (15.53) ^a	53.79% (18.49) ^b	77.37% (2.86) ^{a,c,d}	23.61% (12.00) ^{b,c,e,f}	54.29% (14.55) ^e	47.67% (20.42) ^{d,f}	49.52% (21.41)	9.43	0.001***

Mixed-effect model comparisons control for clinicians seeing multiple patients and whether the patient was new to the clinic. No adjustment for site was included given collinearity between site and country.

Exact N's vary for each row, due to missing data. Data available upon request.

*p<0.05; **p<0.01, ***p<0.001

a,b,c,d,e,f,g,h. Values with paired superscripts in the same row differ significantly (p< 0.05) after adjusting for multiple comparisons, Tukey-Kramer test.

Table DS8 Qualitative data on reasons for feasibility, acceptability, and clinical utility, differentiated by patients and clinicians

Code	Operationalized Sub-code	Magnitude	Patients (N=318)		Representative Quotes	Clinicians (N=318)		Representative Quotes
			n=Total by Distinct Patients	N=Total by Text Coded		n=Total by Distinct Clinicians	N=Total by Text Coded	
Feasibility: Any discussion of how the CFI can be used in service settings.	Issues related to the CFI as a tool	Positive response	81	110	"I thought it was really good. You go from basic questions to more complex. Complex in the sense of how you feel."	30	38	"Having a set of questions clubbed together focuses or brings my attention to the cultural, background aspects of patients. It brings my attention back to these factors, which is definitely good for me in terms of a reminder."
		Neutral/indifferent response	10	13	"I did not understand it initially but then I think I got it."	11	14	"I think it's the kind of thing I would like everybody to be trained on. When they're evaluating someone for the first time, they can use these questions towards the end of the interview to ask anything that was not elicited. I don't think I could see it being used at the beginning of the interview."
		Negative response	26	36	"It was troubling because there were a lot of questions. I don't understand. All the details confuse me."	107	274	"Compared to other diagnostic interviews I've done, I feel things got jumbled."
	Issues related to implementing the CFI within a clinical setting	Positive response	14	17	"The difference is the patience of the doctor. I didn't notice a pressure in him. I didn't feel forced."	9	9	"It can definitely be used at the intake process."
		Neutral/indifferent response	3	3	"It didn't affect anything. I was not comfortable, but there were no major problems."	6	6	"The CFI is mostly relevant for non-language concordant services."
		Negative response	7	9	"I don't know if doctors can spend this much time with patients."	39	47	"Human resources are limited and the extra time required for this will increase patient waiting time which is already strained."
Acceptability -Any discussion of how the CFI elicits emotions among patients and clinicians.	Positive response		187	350	"It was a good flow. It was calm, easy, relaxed. I wasn't stressed. I wasn't nervous. I felt like I was talking to someone I knew, like a friend or something like that."	39	52	"It allows me to empathize more with the patient."

	Neutral/indifferent response		19	20	“It didn’t change my thoughts or feelings.”	4	4	“I feel equally comfortable [addressing cultural aspects of patient presentations] as I was prior to using CFI.”
	Negative response		19	23	“It reminded me of how sad I was and how much I was suffering and worried; it made me think about my future.”	10	11	“I was not at all comfortable. Even though I explained the questions to him [patient], he didn’t get. So I kept wondering what more I could do.”
Clinical Utility <i>-Any discussion of the CFI’s perceived fit, relevance, or compatibility to address a specific clinical problem.</i>	- Diagnosis	Positive response	6	7	“I think it will be better because they will understand what my real mental illness is.”	32	39	“It will help in certain patients who are a diagnostic query. For example, for this patient, initially it seemed like psychosis, but it wasn’t so upon talking to the patient at length.”
		Neutral/ Indifferent response	1	1	“I cannot say. Maybe it helps me. It might help the caretaker to draw conclusions.”	15	15	“It doesn’t seem to modify the diagnosis.”
		Negative response	0	0	n/a	5	5	“It significantly lengthens interview time without getting a clear diagnostic picture.”
	- Treatment	Positive response	137	202	“The CFI will help me get better treatment because it will help the team better assess me and my problem.”	63	80	“What brings them to the clinic is always very important and helps get a general idea, because that sets up what they consider the problem to be and how you can engage that problem with treatment.”
		Neutral/ Indifferent response	15	15	“I don’t think that there will be any effects.”	8	9	“I don’t think it will have any impact on treatment.”
		Negative response	1	1	“The questions couldn’t help me improve myself.”	1	1	“Cultural facts might not help in deciding the pharmacological management.”
	- Role of culture in mental illness	Positive response	21	27	“The doctors need to understand that patients from this country are this way and they are that way from this other country. They are all not the same.”	61	83	“This will help me understand the patient’s problem extensively on the basis of cultural/ religious things.”
		Neutral/ Indifferent response	2	2	“That question about our cultural background was not about rites, rituals, and religion. It didn’t apply to us, so those didn’t seem so helpful.”	15	15	“I’m finding that there are some people, like this particular patient who I saw today, who jump on it and say, ‘This is how culture has affected me,’ and other people who look at me like, ‘Are you crazy? I don’t even know what you’re

								talking about'."
		Negative response	11	11	"I don't think that asking people about their religious backgrounds is right."	23	24	"Different cultures have different beliefs, and to incorporate these might be difficult."
	- General information gathering (not specific to diagnosis or treatment)	Positive response	154	300	"I would say in this interview that today we touched on a lot of things that I would have taken many different sessions to discuss with my talk therapist or psychiatrist. So there's a lot more personal information in a shorter amount of time."	106	159	"I learned how he being able to talk about his symptoms opened up his family to talking about the symptoms."
		Neutral/Indifferent response	24	26	"It seems like kind of the same; the types of interviews I've given in the past are very similar."	10	11	"The CFI provides some psychosocial information, but that is already assessed in a good clinical history taking."
		Negative response	6	6	"The only thing is that sometimes one needs to investigate a little about the problem, about your mental or emotional problem. And those are the things that I don't like to talk about very frequently. I don't like to talk much about the problem."	17	21	"I think they need to streamline [the question about] the groups that have been helpful or not helpful. I think there's almost too much emphasis on outside groups as opposed to what the individual has experienced."

n/a - not applicable