


Measuring Patient Experiences of Integration in Health Care Delivery: Psychometric Validation of IntegRATE Under Controlled Conditions

Journal of Patient Experience
Volume 8: 1-9
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23743735211007346
journals.sagepub.com/home/jpx


Rachel Thompson, PhD^{1,*} , Gabrielle Stevens, PhD^{2,*} ,
and Glyn Elwyn, MB, BCh, PhD² 

Abstract

The objective of this study was to assess the psychometric properties of IntegRATE—a 4-item patient-reported measure of integration in health care delivery—under controlled conditions. Adults who reported having received health care in the previous year were exposed to a fictional health care scenario featuring good, mixed, or poor integration on 1 or 2 occasions. They were then asked to imagine themselves as a patient in the scenario and complete IntegRATE and other measures. The data collected were analyzed to assess the discriminative, concurrent, and divergent validity of IntegRATE and its test–retest reliability and responsiveness using both “sum score” and “top score” scoring approaches. Six-hundred people participated in the study with 190 taking part on 2 occasions. The IntegRATE sum score demonstrated discriminative validity, concurrent validity, divergent validity, and responsiveness and partially demonstrated test–retest reliability. The IntegRATE top score demonstrated concurrent validity, divergent validity, and responsiveness and partially demonstrated discriminative validity and test–retest reliability. We conclude that the IntegRATE sum score exhibits encouraging psychometric properties and performs more optimally than the IntegRATE top score.

Keywords

health literacy, interprofessional communication, measurement, patient feedback, team communication, transitions of care, patient satisfaction

Introduction

Integration in the delivery of health care is “a way of organizing care delivery—by coordinating different activities to ensure harmonious functioning—ultimately to benefit the patients in terms of clinical outcome” (1). Although some indicators of integration can be observed only by those who deliver health care, others are experienced directly by patients themselves. Four salient markers of health care integration from the patient perspective are effective information transfer between health care team members, consistent information provision by team members, respect and collaboration among team members, and clarity in the roles and responsibilities of different team members (2).

There are significant advantages to integration in health care delivery and these have led to elevation of the concept in health policy and clinical practice guidelines around the world. A recent systematic review of 167 studies evaluating models of integrated care, for example, found that

integration improved patients’ satisfaction, perceived quality of care, and access to services (3). This and other evidence suggests that efforts to enhance integration may be particularly beneficial for people with chronic disease, people with multiple health problems, and people underserved by current health services and systems.

Being able to measure the degree of integration in health care delivery from the patient perspective is important.

¹ School of Public Health, Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, Australia

² The Dartmouth Institute for Health Policy & Clinical Practice, Dartmouth College, Hanover, NH, USA

* These authors are joint lead authors of this article.

Corresponding Author:

Rachel Thompson, School of Public Health, Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales 2006, Australia.
Email: rachel.thompson@sydney.edu.au



IntegRATE was developed to address the absence of a valid, reliable, and scalable patient-reported measure of health care integration (2). IntegRATE comprises 4 items that assess patient experiences of health care in the domains of information sharing, consistent advice, mutual respect, and role clarity. Cognitive interviews with members of the public found that integRATE items were clearly understood and a subsequent pilot administration of the measure demonstrated that it could be completed in less than one minute (2). However, the psychometric properties of IntegRATE remained unassessed, precluding insights into whether it can be used to validly and reliably measure health care integration from the patient perspective. The objective of this study was to use simulated health care experiences to examine whether IntegRATE distinguishes between poor, moderate, and good integration in health care delivery (discriminative validity), captures meaningful changes in integration (responsiveness), produces the same results over time (test–retest reliability), and relates to other measures as expected (concurrent validity and divergent validity) (4).

Methodology

Design and Allocation

We conducted a 3×2 mixed fractional factorial study. The between-subjects factor was the degree of integration featured in a simulated health care experience to which participants were exposed (good integration, mixed integration, poor integration) and the within-subjects factor was time (Time 1 (T1), Time 2 (T2)). At T1, participants were randomized to either the good integration, mixed integration, or poor integration condition. At T2, participants were randomized to either the good integration or poor integration condition. All randomization was automated within the online survey platform.

Participants and Recruitment

Participants were recruited by Qualtrics, a commercial panel service based in the United States. Prospective participants were provided with a link to a study information sheet and invited to provide informed electronic consent to participate in a survey. Those who provided consent, were aged 18 years or older, and responded affirmatively when asked, “Have you had a health issue that has led you to see different people (eg, office staff, nurses, doctors, and other health professionals) in the past 12 months?” could proceed immediately to the T1 survey unless a prespecified quota corresponding to their sociodemographic characteristics (ie, age, gender, educational attainment) had already been met. Participants were offered compensation by the panel service, which may have included cash, airline miles, gift cards, redeemable points, sweepstakes entries, or vouchers.

A subset of study participants (ie, those randomized to the “good integration” or “poor integration” conditions during the T1 survey) were recontacted 1 to 3 weeks later. They were again provided with a link to a study information sheet

and invited to provide informed electronic consent to participate in a survey. Those who provided consent could proceed immediately to the T2 survey unless an overall participant quota had already been met. As before, participants were offered compensation by the panel service.

Materials

We created 3 fictional letters written by a couple describing a recent health care experience as well as audio clips of the letters being read aloud. The content of the letters was informed by the conceptualization of health care integration offered by IntegRATE. The letters followed a standard structure, were matched on word length and formatting, and varied only in the degree of integration (ie, good, mixed, or poor) featured in the experience (see Table 1 for key excerpts and Supplementary File for complete letters). Maternity care was chosen as the context for the health care experience because it was considered a relatively relatable episode of care that usually involves several people. The tone and language in the letters was informed by our previous research on patient experiences of maternity care. The letters were also reviewed by a midwife and obstetrician to ensure relevance to the United States context. The letters were assessed as having a Flesch Reading Ease range of 74.3 to 75.1 and a Flesch-Kincaid Grade Level range of 6.8 to 6.9.

Data Collection

Time 1 (T1). Participants reported their age, gender, educational attainment, race and ethnicity, and language(s) spoken at home (5). Participants’ health literacy was also assessed using a single item that asked, “How confident are you filling out medical forms by yourself?” (6,7) In line with recommendations, we classified responses of “Not at all,” “A little bit,” and “Somewhat” as indicative of limited health literacy (8,9) and “Quite a bit” and “Extremely” as indicative of adequate health literacy. Participants were then presented with the letter (and audio clip) corresponding to their randomized condition, were asked to “Please imagine you are one of the patients who wrote the letter and answer the following questions,” and were presented with IntegRATE (2), the Role Clarity and Coordination within Clinic subscale of the Patient-Perceived Continuity of Care from Multiple Clinicians scale (10), and an item assessing perceptions of the hospital’s receptivity to feedback (survey available on request). The approach and language adopted to encourage participants to imagine themselves in the simulated health care experience were informed by previous research (11).

IntegRATE assesses integration in health care delivery using 4 items (see Table 2) with responses assessed on a 3-point scale (3 = “Never,” 2 = “A little,” 1 = “A lot,” 0 = “Always”). For this study, IntegRATE was scored in 2 ways. The IntegRATE sum score (range 0-12) was calculated by summing item responses, with higher values indicating more integration. The IntegRATE top score was

Table 1. Letter Excerpts Demonstrating the Manipulation of Integration Across Conditions.

Good integration condition	Mixed integration condition	Poor integration condition
Domain: Consistent Advice		
When I began to get a headache, we were nervous so we called the hospital. The person on the phone agreed with our doctor. Then, when we told her about my headache, she said this was a good reason to come in to the hospital right away. We were clear it was the right thing to do and chose to come in.	(As for poor integration condition)	When I began to get a headache, we were nervous so we called the hospital. The person on the phone disagreed with our doctor. Then, when we told her about my headache, she said that this was common and to wait to see if it would settle. We were unclear about what to do but chose to come in, just in case.
Domain: Information Sharing		
When we arrived at the maternity unit, the person at the front desk knew we were coming, so we were able to go straight through to the maternity unit. Then, when we were getting settled in our room, our nurse came in. He had been told about my headache, so my blood pressure was taken right away.	(As for poor integration condition)	When we arrived at the maternity unit, the person at the front desk didn't know we were coming, and we were asked the same questions again. Then, when we were getting settled in our room, our nurse came in. He had not been told about my headache, so my blood pressure wasn't taken until 30 minutes later.
Domain: Mutual Respect		
Within the first couple of hours, a doctor, a breastfeeding specialist, and our nurse had all come by to help. They really seemed to enjoy working together, which made a stressful situation so much easier for us to deal with.	(As for good integration condition)	Within the first couple of hours, a doctor, a breastfeeding specialist, and our nurse had all come by to help. They didn't seem to enjoy working together, which made a stressful situation so much harder for us to deal with.
Domain: Role Clarity		
While Sam massaged my lower back, where it was hurting most, our doctor came into the room with 2 new people. Our doctor explained that they were the doctor and nurse who had come to do the epidural. Knowing why everyone was there made us feel more in control.	(As for good integration condition)	While Sam massaged my lower back, where it was hurting most, our doctor came into the room with 2 new people. Our doctor did not explain who these people were or what they were going to do. Not knowing why everyone was there made us feel less in control.

Table 2. IntegRATE Domains and Items.

Domain	Item
Information sharing	How often did you have to do or explain something because people did not share information with each other?
Consistent advice	How often were you confused because people gave you conflicting information or advice?
Mutual respect	How often did you feel uncomfortable because people did not get along with each other?
Role clarity	How often were you unclear whose job it was to deal with a specific question or concern?

calculated by applying a code of 1 (high integration) if the response to all 4 items was "Never" and a code of 0 (limited integration) if a participant's response to 1 item or more was "A little," "A Lot," or "Always."

The 3-item Role Clarity and Coordination within Clinic subscale of the Patient-Perceived Continuity of Care from Multiple Clinicians scale measures the extent to which a patient is given conflicting information and observes clinicians not working well together (10). Minor adaptations were made to item wording to suit the clinical context described in the letters. Participant responses were dichotomized and then summed, yielding a total score (range 0-3) where higher values indicated less role clarity and coordination.

A newly developed item assessed perceptions of the extent to which the hospital featured in the letter was receptive to feedback from patients. Responses were assessed on a

5-point scale (1 = "Not at all interested," 2 = "Slightly interested," 3 = "Moderately interested," 4 = "Very interested," 5 = "Extremely interested").

Time 2 (T2). Participants were presented with the letter (and audio clip) corresponding to their newly randomized condition, were again asked to imagine being one of the patients who wrote the letter, and were asked to complete same 3 instruments as at T1 (survey available on request).

Analysis

Preliminary analyses. Preliminary analyses were conducted to compare the sample of participants to the US adult population on sociodemographic characteristics and to compare the equivalence of participants across the 3 conditions at T1 where cell sizes permitted.

Main analyses. Planned analytic methods for assessing the discriminative validity, concurrent validity, divergent validity, test–retest reliability, and responsiveness of the IntegRATE sum score and the IntegRATE top score (see Supplementary File) were adapted from Barr et al (11).

Planned subgroup analyses. We planned to assess the validity and reliability of the IntegRATE sum score and the IntegRATE top score among subgroups based on gender and health literacy where sample sizes permitted. Ultimately, subgroup analyses were only conducted for the IntegRATE sum score due to concerns about statistical power for the IntegRATE top score and the established superiority of the IntegRATE sum score in the main analyses. Other subgroup analyses that were not reported due to small sample sizes are indicated.

Data quality and treatment of missing data. We excluded participants who completed the T1 survey in less than half of the median completion time (based on the first 10% of participants) from all study analyses. We excluded any participant who completed T2 survey in less than half of the median time (based on all participants) from all T2 analyses only (ie, test–retest reliability and divergent validity). Participants with any missing data on IntegRATE at T1 were excluded from all analyses and participants with any missing data on IntegRATE at T2 were excluded from all T2 analyses only. Participants with missing data on the Role Clarity and Coordination within Clinic subscale and/or the item assessing hospital receptivity to feedback were excluded from the relevant analyses only. Participants who did not report a “Female” or “Male” gender ($n = 4$) were not included in subanalyses based on gender. Participants with missing data on health literacy ($n = 5$), all of whom reported a highest level of schooling as “No schooling completed, or less than 1 year” or “Nursery, kindergarten, and elementary (grades 1-8)” were coded as having limited health literacy.

Sample Size

In the absence of research on which to base effect size estimates, we chose a target sample size of 50 participants per cell for the analyses that used data collected on 2 occasions (ie, test–retest reliability and responsiveness). Assuming that 50% of eligible participants would participate in the study on both occasions, we had an overall target sample size of 600 participants.

Results

Participants

Six hundred and seventy-four participants were randomized to a study condition at T1 and 600 were included in analyses. Four hundred and four participants were invited to participate in the T2 survey and 190 were included in

analyses (see Figure 1). The average time elapsed between the T1 and T2 surveys was 16.6 days (SD = 4.5 days; range = 9.4-27.0 days).

The T1 sample was comparable to the US adult population on gender, age, and educational attainment (12). The sample overrepresented those not of Hispanic, Latino, or Spanish origin ($P < .001$), those who spoke only English at home ($P < .001$), and those of White only race ($P < .001$) (see Table 3). We found no differences in participant characteristics across the 3 conditions (data available on request).

Discriminative Validity

The discriminative validity of the IntegRATE sum score was demonstrated as there were significant differences in scores between participants exposed to good and mixed integration scenarios at T1 and between participants exposed to mixed and poor integration scenarios at T1 (see Table 4). The discriminative validity of the IntegRATE sum score was also partially demonstrated for participants with limited health literacy (see Table S1) and demonstrated for participants with adequate health literacy, who identified as female, and who identified as male (see Tables S2-S4).

The discriminative validity of the IntegRATE top score was partially demonstrated. There was a significant difference in the proportion of participants reporting high integration between those exposed to good and mixed integration scenarios at T1 but not between those exposed to mixed and poor integration scenarios at T1 (see Table 4).

Concurrent Validity

The concurrent validity of the IntegRATE sum score was demonstrated by its strong association with the Role Clarity and Coordination within Clinic subscale at T1 (see Table 4). The concurrent validity of the IntegRATE sum score was also demonstrated for all subgroups (see Tables S1-S4).

The concurrent validity of the IntegRATE top score was demonstrated by its moderate association with the Role Clarity and Coordination within Clinic subscale at T1 (see Table 4).

Divergent Validity

The divergent validity of the IntegRATE sum score was demonstrated by its weak association with the item assessing hospital receptivity to feedback at T1 (see Table 4). The divergent validity of the IntegRATE sum score was also demonstrated for all subgroups (see Tables S1-S4).

The divergent validity of the IntegRATE top score was demonstrated by its weak association with the item assessing hospital receptivity to feedback at T1 (see Table 4).

Test–Retest Reliability

The test–retest reliability of the IntegRATE sum score was partially demonstrated. Among all participants exposed to

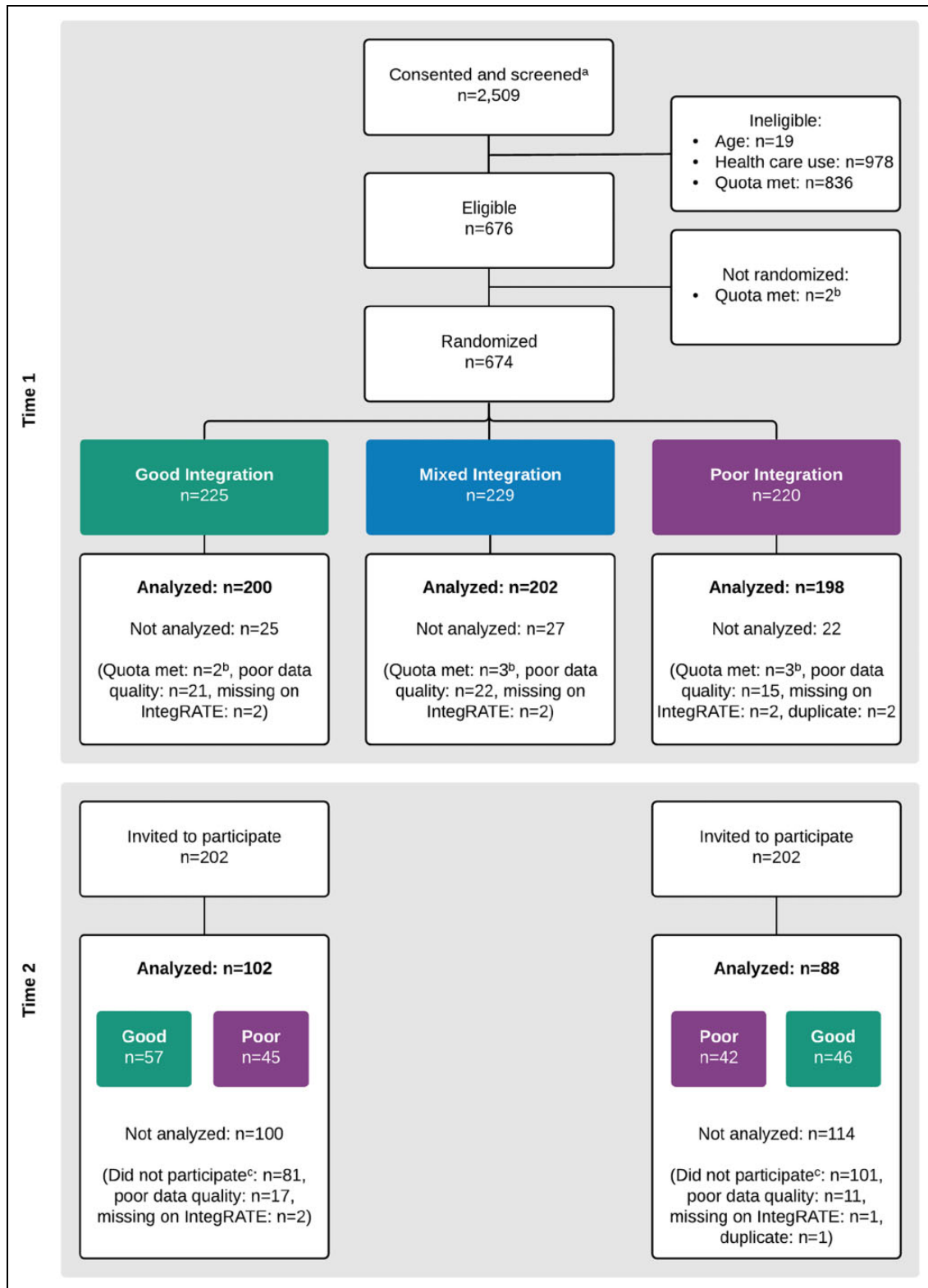


Figure 1. Participant flow diagram. ^aDoes not include those who consented and were screened for the study but discontinued participation. The number of these participants is unknown but is not greater than $n = 454$ given the known number of people who clicked on the survey link. ^bSome participants were excluded because the relevant sociodemographic quota was met, while they were completing the survey. ^cIncludes participants who attempted to take part but were unable to because the participant quota had been met and may also include participants who started the time 2 survey but discontinued for their own reasons. The number of these latter participants is unknown but is not greater than $n = 42$ given the known number of people who clicked on the survey link.

Table 3. Participant Sociodemographic Characteristics by Time 1 Condition.^a

	Good (n = 200) Freq. (%)	Mixed (n = 202) Freq. (%)	Poor (n = 198) Freq. (%)	Total (n = 600) (%)	Population (%)
Gender					
Female	104 (52.0)	98 (48.5)	106 (53.5)	51.3	51.4
Male	95 (47.5)	102 (50.5)	91 (46.0)	48.0	48.6
Other	1 (0.5)	2 (1.0)	1 (0.5)	0.7	-
Age					
18-44 years	98 (49.0)	97 (48.0)	86 (43.4)	46.8	46.8
45-64 years	64 (32.0)	66 (32.7)	76 (38.4)	34.3	33.9
65+ years	38 (19.0)	39 (19.3)	36 (18.2)	18.8	19.3
Educational attainment					
High school graduate or less	77 (38.5)	80 (39.6)	76 (38.4)	38.8	40.9
Some college, no degree, or associate's degree	60 (30.0)	70 (34.7)	71 (35.9)	33.5	31.1
Bachelor's degree or more	62 (31.0)	51 (25.2)	50 (25.3)	27.2	28.0
Prefer not to say	1 (0.5)	1 (0.5)	1 (0.5)	0.5	-
Ethnicity					
Hispanic, Latino, or Spanish origin	17 (8.6)	18 (9.0)	9 (4.6)	7.4	15.5
Not of Hispanic, Latino, or Spanish origin	181 (91.4)	183 (91.0)	186 (95.4)	92.6	84.5
Race					
One race	193 (97.5)	196 (97.5)	188 (95.9)	97.0	97.8
– White	163 (82.3)	165 (82.1)	160 (81.6)	82.0	74.8
– Black or African American	22 (11.1)	19 (9.5)	23 (11.7)	10.8	12.2
– American Indian or Alaska Native	2 (1.0)	1 (0.5)	2 (1.0)	0.8	0.8
– Asian	4 (2.0)	6 (3.0)	0 (0.0)	1.7	5.6
– Native Hawaiian or Other Pacific Islander ^b	0 (0.0)	0 (0.0)	0 (0.0)	0.0	0.2
– Other	2 (1.0)	5 (2.5)	3 (1.5)	1.7	4.4
Two or more races	5 (2.5)	5 (2.5)	8 (4.1)	3.0	2.2
Language spoken at home					
English only	178 (91.3)	180 (91.4)	181 (93.8)	92.1	78.7
Language(s) other than English	17 (8.7)	17 (8.6)	12 (6.2)	7.9	21.3
Health literacy					
Limited	27 (13.5)	46 (22.8)	35 (17.7)	18.0	^c
Adequate	173 (86.5)	156 (77.2)	163 (82.3)	82.0	^c

^aFrequencies may not add to the total due to occasional cases of missing data.

^bNot included in population comparison due to cell count of 0.

^cNo population data are available.

the same integration scenario at T1 and T2, there was good reliability in scores over time. For participants exposed to the good integration scenario at T1 and at T2, there was moderate reliability in scores over time, but for participants exposed to the poor integration scenario at T1 and T2, reliability was poor (see Table 4). The test–retest reliability of the IntegRATE sum score was also partially demonstrated for participants with adequate health literacy, who identified as female, and who identified as male (see Tables S2-S4) but could not be determined for participants with limited health literacy (see Table S1).

The test–retest reliability of the IntegRATE top score was partially demonstrated. Among all participants exposed to the same integration scenario at T1 and T2, there was moderate agreement in scores over time. However, agreement was fair for participants exposed to the good integration scenario at T1 and T2 and could not be determined for those exposed to the poor integration scenario at T1 and T2 (see Table 4).

Responsiveness

The responsiveness of the IntegRATE sum score was demonstrated by significant differences in scores when participants were assigned to the good integration scenario and then the poor integration scenario, and when participants were assigned to the poor integration scenario and then the good integration scenario (see Table 4). The responsiveness of the IntegRATE sum score was also demonstrated for participants with adequate health literacy, who identified as female, and who identified as male (see Tables S2-S4) but could not be determined for participants with limited health literacy (see Table S1).

The responsiveness of the IntegRATE top score was demonstrated by significant differences in the proportion of participants reporting high integration when they were exposed to the good integration scenario and then the poor integration scenario, and when they were exposed to the poor integration scenario and then the good integration scenario (see Table 4).

Table 4. Validity and Reliability of the IntegRATE Sum Score and Top Score: All Participants.

Property	Analysis	N	Result	Interpretation	Demonstrated
Sum score					
Discriminative validity	Difference between good (M = 9.68, SD = 2.87) and mixed (M = 7.06, SD = 2.52) integration conditions at time 1	402	MD = 2.61, $p < .001$	Significant difference	Yes
	Difference between mixed (M = 7.06, SD = 2.52) and poor (M = 5.14, SD = 2.72) integration conditions at time 1	400	MD = 1.93, $p < .001$	Significant difference	Yes
Concurrent validity	Association between IntegRATE score and Role Clarity and Coordination within Clinic subscale score at time 1	597	$r = -0.75$, $p < .001$	Strong-negative association	Yes
Divergent validity	Association between IntegRATE score and hospital receptivity to feedback score at time 1	600	$r = 0.24$, $p < .001$	Weak-positive association	Yes
Test–retest reliability	Agreement between time 1 and time 2 scores among all randomized to the same integration condition	99	ICC (3,1) = 0.78, $p < .001$	Good reliability	Yes
	Agreement between time 1 and time 2 scores among those randomized to the good integration condition	57	ICC (3,1) = 0.68, $p < .001$	Moderate reliability	Yes
	Agreement between Time 1 and Time 2 scores among those randomized to the poor integration condition	42	ICC (3,1) = 0.32, $p = .020$	Poor reliability	No
Responsiveness	Difference between good integration condition at time 1 (M = 10.07, SD = 2.60) and poor integration condition at time 2 (M = 4.91, SD = 3.29)	45	MD = 5.16, $p < .001$	Significant difference	Yes
	Difference between poor integration condition at time 1 (M = 4.78, SD = 2.29) and good integration condition at time 2 (M = 9.93, SD = 2.64)	46	MD = -5.15, $p < .001$	Significant difference	Yes
Top score					
Discriminative validity	Difference between good (39.5% high integration) and mixed (2.5% high integration) integration conditions at time 1	402	$\chi^2 (1) = 83.35$, $p < .001$	Significant difference	Yes
	Difference between mixed (2.5% high integration) and poor (3.5% high integration) integration conditions at time 1	400	$\chi^2 (1) = 0.39$, $p = .534$	No difference	No
Concurrent validity	Association between IntegRATE score and Role Clarity and Coordination within Clinic subscale score at time 1	597	$r_{pb} = -0.49$, $p < .001$	Moderate negative association	Yes
Divergent validity	Association between IntegRATE score and hospital receptivity to feedback score at Time 1	600	$r_{pb} = 0.23$, $p < .001$	Weak-positive association	Yes
Test–retest reliability	Agreement between time 1 and time 2 scores among all randomized to the same integration condition	99	Agreement = 76.8; $\kappa = 0.43$, $p < .001$	Moderate agreement	Yes
	Agreement between time 1 and time 2 scores among those randomized to the good integration condition	57	Agreement = 63.2; $\kappa = 0.30$, $p = .013$	Fair agreement	No
	Agreement between time 1 and time 2 scores among those randomized to the poor integration condition	42 ^a			Unknown
Responsiveness	Difference between good integration condition at time 1 (40.0% high integration) and poor integration condition at Time 2 (4.4% high integration)	45	$p < .001$	Significant difference	Yes
	Difference between poor integration condition at time 1 (0% high integration) and good integration condition at time 2 (50.0% high integration)	46	$p < .001$	Significant difference	Yes

Abbreviations: M, mean; SD, standard deviation; MD, mean difference; ICC, intraclass correlation coefficient.

^a Results not reported due to extremely low cell numbers.

Discussion

In this assessment of IntegRATE under controlled conditions, the IntegRATE sum score demonstrated encouraging psychometric properties. It yielded incrementally higher scores as the degree of integration to which participants were exposed increased, whether analyses were conducted between participants (discriminative validity) or within participants over time (responsiveness). It also yielded scores that were strongly correlated with the 3-item Role Clarity and Coordination within Clinic subscale of the Patient-Perceived Continuity of Care from Multiple Clinicians scale (10) (concurrent validity) and weakly correlated with a new item assessing perceptions of hospital receptivity to patient feedback (divergent validity). Repeated administration of IntegRATE over time also yielded IntegRATE sum scores that were correlated (test–retest reliability), but further analyses indicated that this was evident only for participants exposed to good integration and not for participants exposed to poor integration. In contrast, the IntegRATE top score performed relatively poorly in this study, only partially demonstrating discriminative validity.

For the IntegRATE sum score, subgroup analyses conducted among participants with limited and adequate health literacy and among female and male participants generated largely the same conclusions as for the sample as whole. Particularly encouraging was the demonstration of concurrent validity and divergent validity and partial demonstration of discriminative validity among participants with limited health literacy. However, because we did not power this study for subgroup analyses and there was only a small sample size for some analyses, we recommend that these subgroup findings be interpreted with caution and explored further in future studies.

One of the main strengths of this study was our use of simulated health care experiences that enabled us to test psychometric hypotheses in a systematic and controlled way. Many of the conclusions that we were able to draw from this study, particularly those pertaining to the discriminative validity of IntegRATE, would not otherwise have been possible without the allocation of significant resources to the independent assessment of the integration present in an entire episode of patient care. Efforts to maximize data quality, including the use of demographic quotas to enhance population representativeness, the development of fictional letters informed by our previous patient experience research and provision of an audio version of each letter to enhance understanding by participants with lower levels of literacy comprise further strengths of this study. The principal study limitation relates to the unknown generalizability of assessments of integration made in response to a concise written and audio account of an episode of care as opposed to one experienced firsthand (potentially over an extended period of time). A second study limitation was our inability to conduct some planned subgroup analyses.

Conclusion

We conclude that IntegRATE, when scored using the sum score approach, is a promising patient-reported instrument for assessing integration in the delivery of health care. Given that its brevity, simplicity, and condition-neutral focus make IntegRATE relatively feasible for adoption in routine practice, particularly as compared to other measures (2), we recommend research confirming its psychometric properties in the clinical setting, including when administered among diverse patient populations. Related work assessing the minimum clinically meaningful difference as assessed by IntegRATE is also warranted.

Authors' Note

Ethical approval for this study was obtained from the Dartmouth College Committee for the Protection of Human Subjects (#29038). Electronic informed consent was obtained from study participants for their anonymized data to be published in this article. All procedures in this study were conducted in accordance with protocols approved by the Dartmouth College Committee for the Protection of Human Subjects.

Acknowledgments

We are grateful to Dr Kyla Donnelly for assisting in the drafting of the letters and development of the online survey and to Dr Shama Alam for voicing the audio clips used in the study.

Author Contributions

RT conceived and designed the study and drafted the manuscript. GS was responsible for data analysis and contributed to drafting of the manuscript. GE contributed to the conception and design of the study and revision of the manuscript.




Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: RT and GE are owners of copyright in IntegRATE. RT and GE have not participated in any efforts to exploit this copyright commercially and have not received any personal income connected to this ownership. GS and GE are named investigators on funding awarded to Dartmouth College for research projects that use and study IntegRATE.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Recruitment and data collection costs were met by internal funds available to RT and GE.

ORCID iDs

Rachel Thompson  <https://orcid.org/0000-0002-1009-0746>
Gabrielle Stevens  <https://orcid.org/0000-0001-9001-178X>
Glyn Elwyn  <https://orcid.org/0000-0002-0917-6286>

Supplemental Material

Supplemental material for this article is available online.

References

1. Strandberg-Larsen M, Krasnik A. Measurement of integrated healthcare delivery: a systematic review of methods and future research directions. *Int J Integr Care*. 2009;9:e01.
2. Elwyn G, Thompson R, John R, Grande SW. Developing IntegrATE: a fast and frugal patient-reported measure of integration in health care delivery. *Int J Integr Care*. 2015;15.
3. Baxter S, Johnson M, Chambers D, Sutton A, Goyder E, Booth A. The effects of integrated care: a systematic review of UK and international evidence. *BMC Health Serv Res*. 2018;18:350.
4. Streiner D, Norman G. *Health Measurement Scales: A Practical Guide to Their Development and Use*. Oxford University Press; 2008.
5. US Census Bureau. *The American Community Survey 2015*. 2015. Washington, DC: US Census Bureau.
6. Dageforde LA, Cavanaugh KL, Moore DE, Harms K, Wright A, Pinson CW, et al. Validation of the written administration of the Short Literacy Survey. *J Health Commun*. 2015;20:835-842.
7. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. *Fam Med*. 2004;36:588-594.
8. Chew LD, Griffin JM, Partin MR, Noorbaloochi S, Grill JP, Snyder A, et al. Validation of screening questions for limited health literacy in a large VA outpatient population. *J Gen Intern Med*. 2008;23:561-566.
9. Wallace LS, Rogers ES, Roskos SE, Holiday DB, Weiss BD. Screening items to identify patients with limited health literacy skills. *J Gen Intern Med*. 2006;21:874-877.
10. Haggerty JL, Roberge D, Freeman GK, Beaulieu C, Bréton M. Validation of a generic measure of continuity of care: when patients encounter several clinicians. *Ann Fam Med*. 2012;10:443-450.
11. Barr PJ, Thompson R, Walsh T, Grande SW, Ozanne EM, Elwyn G. The psychometric properties of CollaboRATE: a fast and frugal patient-reported measure of the shared decision-making process. *J Med Internet Res*. 2014;16.
12. US Census Bureau. 2015 American Community Survey 1-year estimates [Internet]. Accessed September 26, 2016. <https://data.census.gov/cedsci/all?q=UnitedStates&y=2015>

Author Biographies

Rachel Thompson is a senior research fellow at the School of Public Health, University of Sydney, Australia.

Gabrielle Stevens is a research scientist at The Dartmouth Institute for Health Policy & Clinical Practice, Dartmouth College, USA.

Glyn Elwyn is a professor at The Dartmouth Institute for Health Policy & Clinical Practice, Dartmouth College, USA.