



Interrater reliability for the Four Habits Coding Scheme as part of a randomized controlled trial

Bård Fossli Jensen^{a,*}, Pål Gulbrandsen^{a,b}, Jurate Saltyte Benth^{a,b}, Fredrik A. Dahl^a, Edward Krupat^c, Arnstein Finset^d

^a HØKH, Research Centre, Akershus University Hospital, Lørenskog, Norway

^b Faculty Division, Akershus University Hospital, University of Oslo, Oslo, Norway

^c Center for Evaluation, Harvard Medical School, Boston, MA, USA

^d Department of Behavioural Sciences, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway

ARTICLE INFO

Article history:

Received 1 March 2010

Received in revised form 24 June 2010

Accepted 25 June 2010

Keywords:

Interrater reliability

Randomized controlled trial

RCT

IRR

Four Habits Coding Scheme (4HCS)

ABSTRACT

Objective: To describe the process for developing interrater reliability (IRR) for the Four Habits Coding Scheme (4HCS) for a heterogeneous material as part of a randomized controlled trial.

Methods: Videotapes from 497 hospital encounters involving 71 doctors from most clinical specialties were collected. Four experienced psychology students were trained as raters. We calculated Pearson's *r* and the intraclass correlation (ICC) on the total score across consecutive samples of twenty videos, and Pearson's *r* on single videos across items in the initial coding phase.

Results: After 18 h of training and one rating session, the total score Pearson's *r* and ICC exceeded .70 for all pairs of raters. Across items within single videos, the Pearson's *r* was never below 0.60 after the first 50 videos. At item and habit level Pearson's *r* remained unsatisfactory for some rater pairs mostly due to low variance on some items.

Conclusion: Based on the evaluation of the effect of communication skills training via a total score, IRR was satisfactory for the 4HCS as applied to heterogeneous material. However, good reliability at item level was difficult to achieve.

Practice implications: 4HCS may be used as an outcome measure for clinical communication skills in randomized controlled trials.

© 2010 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

When evaluating doctor behavior for research purposes using videotapes, a satisfactory degree of agreement among raters, commonly known as interrater reliability (IRR) [1], must be established. Many ways can be used to approach this task. The choice of method is influenced by the following three main factors: (1) the purpose of the study, (2) the nature of the data, and (3) the available time and resources [2]. The purpose of our study was to evaluate the effect of a communication skills training program across specialties in a hospital. The sum scores of 23 items on 497 videotapes were collected. Since the timeframe for this study was limited, the use of more than two raters was necessary. This paper describes how we established IRR under these circumstances.

Interventions aimed at improving communication skills are quite common [3]. In order to assess the effects of such interventions, randomized controlled trials (RCT) should ideally be performed. The RCT paradigm is quantitative by nature, and a substantial number of observations are often needed. A valid measure for the evaluation of communication skills must be used, and it should be reasonably easy to apply. Preferably, this measure should be a score that reliably assesses the concept of good communication as derived from observation, e.g., behavior in video or audio taped consultations. Unfortunately, few tools that satisfy these needs for validity and ease of collection have been developed to date [4,5]. One recent approach, the Four Habits Coding Scheme (4HCS) (Fig. 1), is a coding scale that was constructed and validated [6] to assess the effect of the “Four Habits approach to effective clinical communication” (The Four Habits Model) as developed by Kaiser Permanente [7]. The Four Habits Model is a training program that organizes well-known communication principles into subgroups (i.e., habits) for didactic purposes, thereby making them easy to teach and remember. The four habits are as follows: invest in the beginning to create rapport and set an agenda (i.e.,

* Corresponding author at: Akershus University Hospital, HØKH, Research Centre, Sykehusveien 27, Nordbyhagen, NO 1474, Norway. Tel.: +47 22851020; fax: +47 22851300.

E-mail address: b.f.jensen@medisin.uio.no (B. Fossli Jensen).

Four Habits Coding Scheme

Code each of the items below using the categories 1, 3 or 5. If you feel strongly that the behaviour being coded is directly between categories, you may use the categories 2 or 4.

	Not very Effective			Highly Effective	
1. Invest in the Beginning					
A. Shows familiarity with patient	1	(2)	3	(4)	5
B. Greets patient warmly	1	(2)	3	(4)	5
C. Makes small talk	1	(2)	3	(4)	5
D. Uses primarily open-ended questions	1	(2)	3	(4)	5
E. Encourages expansion of patient's concerns	1	(2)	3	(4)	5
F. Elicits the full agenda of concerns	1	(2)	3	(4)	5
2. Elicit the patient's Perspective					
A. Interested in patient's understanding of problem	1	(2)	3	(4)	5
B. Asks about patient's goal for visit	1	(2)	3	(4)	5
C. Shows interest in impact on patient's life	1	(2)	3	(4)	5
3. Demonstrate Empathy					
A. Encourages expression of emotions	1	(2)	3	(4)	5
B. Accepts / validates patient's feelings	1	(2)	3	(4)	5
C. Helps to identify / label feelings	1	(2)	3	(4)	5
D. Displays effective nonverbal behaviour	1	(2)	3	(4)	5
4. Invest in the End					
A. Frames information using patient's perspective	1	(2)	3	(4)	5
B. Allows time for information to be absorbed	1	(2)	3	(4)	5
C. Explains clearly / uses little jargon	1	(2)	3	(4)	5
D. Explains rationale for tests and treatments	1	(2)	3	(4)	5
E. Effectively tests for comprehension	1	(2)	3	(4)	5
F. Encourages involvement in decision-making	1	(2)	3	(4)	5
G. Explores acceptability of treatment plan	1	(2)	3	(4)	5
H. Explores barriers of implementation	1	(2)	3	(4)	5
I. Encourages additional questions	1	(2)	3	(4)	5
J. Makes clear plans for follow-up	1	(2)	3	(4)	5

Fig. 1. Four Habits Coding Scheme.

Habit I), elicit the patient's perspective (i.e., Habit II), demonstrate empathy in a constructive way (i.e., Habit III), and invest in the end to provide information and closure (i.e., Habit IV). Specifically, The Four Habits Model was used in a training program within a teaching hospital in Norway.

In the original study by Krupat et al., the 4HCS was tested through a study of 100 encounters in Boston, MA. The overall IRR in that study was .72 after 8–10 h of training as measured using Pearson's product moment correlation (i.e., Pearson's r) [6]. Validation was accomplished using the well known and much used Roter Interaction Analysis System (RIAS) [6,8]. A 4HCS codebook was developed to describe the qualities in communication that correspond to each value on the scale. Since the 4HCS required no more than 2–5 min over the actual duration of the evaluated consultation [6] according to scholarly reports and was based on The Four Habits Model, which was already validated, we decided to apply the 4HCS to assess the effects of the training program.

IRR is more accurately defined as the level of agreement among a specific set of raters on a specific instrument at a specific time. It is a property of the testing situation, not of the tool itself, and the value has important implications for the validity of the study results [9]. Although the original study of Krupat et al. reported sufficient IRR, that study used a homogeneous sample of primary care consultations, and only two raters performed the coding. Our study was different in significant ways: specifically, the following are the important attributes of our study: using a complex design (i.e., RCT), using a hospital setting, occurring on a different continent, involving several specialties, and using more than two raters. Hence, the establishment of IRR was an integral part of the study. The aim of this paper is to describe the rationale for our choices when calculating IRR, share the results of our efforts, and inform future researchers on ways to address IRR when conducting similar studies.

Download English Version:

<https://daneshyari.com/en/article/3815411>

Download Persian Version:

<https://daneshyari.com/article/3815411>

[Daneshyari.com](https://daneshyari.com)