Development and inter-rater reliability of a tool assessing hypnotic communication behaviours adopted by nurses caring for children with cancer: The Sainte-Justine Hypnotic Communication Assessment Scale

Jennifer Aramidéha,b, Terry Mizrahi,a Marie-Claude Charesta, Caroline Plantea, Michel Duvala,c, Serge Sultanb,c,*

a CHU Sainte-Justine, 3175, Chemin de la Côte-Sainte-Catherine, Montréal, Québec, H3T 1C5, Canada
b Department of Psychology, Université de Montréal, Pavillon Marie-Victorin PO Box. 6128, Succursale Centre-ville, Montréal, Québec, H3C 3J7, Canada
c Department of Pediatrics, Université de Montréal, 3175, Chemin de la Côte-Sainte-Catherine, Montréal, Québec, H3T 1C5, Canada

ARTICLE INFO

Keywords:
Hypnotic communication
Healthcare professionals
Assessment tool
Pediatrics
Medical procedures
Procedural pain and distress

ABSTRACT

Background: Several studies in pediatric oncology have shown the successful effects of using hypnotic communication techniques (HCTech) during painful medical procedures. Since no studies assessed the precise use of these techniques with a validated tool, it is unsure that the observed relationships involve the use of HCTech.

Objectives: To develop a scale evaluating healthcare professionals’ behaviours when using HCTech and to evaluate its inter-rater reliability.

Methods: This study involved the preliminary steps of the Sainte-Justine Hypnotic Communication Assessment Scale (SJ-HCAS) development process. As part of a larger intervention study, the SJ-HCAS was developed in three steps by five experts and four lay raters using an iterative process applied to subsets of video-recorded nurse-patient interactions. The development aimed to maximize clarity and precision of items as well as minimize redundancy amongst items. Inter-rater reliability was assessed in a randomly selected sample of 1/3 of collected video-recorded interactions (n = 42).

Results: The final version of the scale is composed of 11 items categorized in two domains pertaining to Relationship and Technique. We found excellent inter-rater reliability for both subscores and total score in two independent inter-rater comparisons (median ICC = 0.879), with most items showing very good to perfect inter-rater reliability (median Kappa = 0.847).

Conclusions: The results support further work with the SJ-HCAS. The scale has the potential to help ensure the integrity of hypnotic communication training in children which could ultimately promote the dissemination of the practice of HCTech.

1. Introduction

Along their cancer trajectory, children have to undergo many painful invasive medical procedures which may affect them daily. It is well-known that children are greatly affected by different types of pain, including pain caused by medical procedures. Among them, venipunctures are a common source of pain in hospitalized children. These needle procedures are associated with a significant level of pain and distress for pediatric patients. In fact, patients report that painful procedures represent one of the most difficult parts of cancer suffering. Compared to adults, procedural pain is a greater source of anxiety and discomfort for children. Studies have even found long and lasting negative consequences years after the end of treatment. Interestingly, previous medical experiences have been shown to be predictive of the child’s reaction to future medical procedures.

It is thus of major importance for healthcare professionals to use appropriate pain and distress management techniques. Different methods have been developed, including pharmacological interventions and non-pharmacological interventions. In pediatric settings, hypnosis is often used to relieve physical or emotional suffering. Clinical hypnosis and hypnosis-derived communication techniques such as hypnoanalgesia (hypnotic suggestions to relieve pain) have great potential as children are easily absorbed in fantasy and imagination. In the Ericksonian tradition, the induction of hypnotic states and phenomena appears to be primarily approached as a matter of communication of ideas and the elicitation of trains of thought and associations.
within the subject and consequent behavioural responses. Moreover, several studies in neuroscience have shown that hypnosis modifies brain activity in the anterior cingulate cortex, which plays an important role in pain modulation.18,19

Previous studies in pediatric oncology have shown that HCTech not only decrease procedure related pain5,20–29 and distress5,20,23–27 but also reduce anxiety5,20,21,23–29 and fear22. However, all these reports call for an independent professional practise hypnosis while the medical procedure is being performed by another healthcare professional (e.g. nurse). None of the studies address the effect of hypnosis communication as used by the professionals (e.g. nurse) themselves. This is a strong limitation to the dissemination of the intervention as it increases costs and is often not feasible in the daily activity of an outpatient clinic. Interestingly, none of the previous studies actually measures treatment integrity, i.e. to what extent the used communication techniques were effectively altered by training and if professionals use HCTech. Consequently, there is no guarantee that the observed relationships (e.g. pre-post differences on a pain scale) actually involve hypnotic communication. In addition, we do not know which components and what intensity in such communication could optimize improvements.

The first objective of this study was to develop a scale assessing pediatric healthcare professionals’ behaviours when using HCTech. The second objective was to evaluate the inter-rater reliability (IRR) of the communication scale. We focused on the level of agreement between raters on scores derived from the scale as well as on individual items.

2. Methods

The scale was developed as part of a research project taking place in our cancer care centre (CHU Sainte-Justine) aiming at evaluating the effects of training nurses to use HCTech in clinical practice (see study protocol31). The purpose of the present scale is to assess pediatric oncology nurses’ behaviours when using HCTech to deal with patients’ procedural pain and distress during venipunctures.

2.1. Participants

To evaluate verbal and nonverbal communication during nurse-patient interactions, venipuncture procedures performed at the CHU Sainte-Justine daycare hematology-oncology clinic were video-recorded. Participant recruitment was completed by May 2015. Six female oncology nurses took part in the study. Inclusion criteria for the nurses were (1) working in the daycare clinic, (2) having experience performing venipunctures with children and (3) having no prior experience in hypnosis. Six patients were assigned to each nurse, using a convenience sampling method.

Eligible patients were identified from the clinic’s computer database. Patients’ inclusion criteria involved (1) being aged between 5 and 18 years old, (2) having a good understanding of French and (3) having
expected regular follow-up at the daycare hematology-oncology clinic. Patients’ exclusion criteria were (1) prior use of hypnoanalgesia, (2) having a psychiatric disorder diagnosis and (3) coming for an emergency or an unscheduled appointment. The first six identified patients for each nurse who met the criteria were then contacted by phone. On the day of their first appointment, patients and their legal guardian met with a research assistant for a short interview (10–15 min) to obtain detailed information about the study as well as consent. A total of thirty-six patients were approached to take part in the study. Of these, three children declined participation because of a lack of interest or not wanting to be exposed to hypnoanalgesia. The final sample for the study consisted of 6 pediatric oncology nurses (6 women, aged: 33 ± 6 yrs) and 33 of their cancer patients (16 boys, 17 girls, aged 10 ± 4 yrs). During the study period, 1 nurse went on maternity leave and 5 patients dropped out of the study. Two patients passed away. Hence, a total of 5 nurses and 26 patients completed the study. All interactions were video-recorded: 117 interactions across the 4-time points (pre- and 2 post-training) were available to evaluate the use of HCTech. Patients as well as nurses provided written informed consent. The study received ethical approval by the CHU Sainte-Justine Research Ethics Committee.

2.2. Development process of the scale

The Sainte-Justine Hypnotic Communication Assessment Scale (SJ-HCAS) was developed by a multidisciplinary team composed of 2 physicians, 1 nurse, 2 psychologists, and 1 psychology graduate student. We used an iterative process aiming to maximize clarity and precision, as well as agreement between raters without redundancy amongst items. The chart summarizing the developing steps is available in Fig. 1.

2.2.1. Objective 1: development of the SJ-HCAS

2.2.1.1. Version 1

First, we conducted an extended literature review to identify important components of hypnotic communication. We collected domains and topics to be covered with researchers involved in the project. This included an in-depth interview with MCC (psychologist and hypnotherapist) and CP (nurse) who conceptualized the training designed for nurses. The training included key elements of the practice of hypnoanalgesia to cover both relational and technical aspects. Key behaviours assessed by the scale were selected from two sources providing details on the practice of hypnosis with children32,33 and a reference guide of hypnotic suggestions34, which also were the references used to design the nurses’ training. Topics and domains were identified by the team, in accordance with the assumption that the practice of basic hypnotic communication requires both establishing a good rapport and using hypnotic techniques relevant to the child’s age and preferences.33

The second step focused on generating a list of items, in which some evaluated relationship abilities (or difficulties) and others the use of (or difficulty with) hypnotic communication techniques. When using hypnotic communication, adapting the language to the client, introducing a slower pace, closely adjusting to the child’s rhythm, developing a cooperative relationship as well as allowing patients to position themselves freely during punctures are all elements reinforcing a trusting nurse-patient relationship. This reinforced relationship will favour the effectiveness of hypnotic suggestions.34 Moreover, using a multisensory stimulation and validating the child’s experience is also common in pediatrics and has been shown to allow a deepening of hypnotic induction.32 Additionally, healthcare professionals focusing their attention on the child as well as using comforting language adapted to the child makes it possible to improve this client-centered approach and individualize the use of hypnotic techniques.32 As changes in children’s behaviours are related to hypnotic communication style, healthcare professionals’ abilities to use adequate techniques and create a hypnotic bubble are pivotal. For each item, one or two examples of behaviours were elaborated to illustrate typical behaviours exemplifying hypnotic communication. To ensure maximum clarity, once each item and example had been chosen by the lead researchers (MCC and SS), we refined the wording by a set of common team discussions (TM, CP, MCC, MD). Clarity was also tested within the team. Following these steps, the Version 1 of the scale was finalized (N = 10 items).

In order to evaluate the clarity of Version 1, a pre-test was conducted. Ten nurse-patient interactions were randomly selected across the 4-time points (pre- and post-training) and analyzed by two raters (psychology graduate students, JA and MPB, see acknowledgements) who did not have prior experience with hypnosis at that time. Raters qualitatively reported on the clarity and face validity of the scales items. Using the same sample of interactions, we compared how raters understood each item and if they were easy to rate. Modifications were expected following this process. The pre-test shed light on several issues with the first version. Raters mentioned having difficulty assigning scores because some of the items were designed to assess more than one behaviour and examples were not sufficiently explicit. Moreover, one difficulty with this version was that items could assess both positive (desirable) and negative (non-desirable) behaviours. This explained why disagreements often occurred between raters, considering one would focus more on desirable behaviours and the other on non-desirable ones.

2.2.1.2. Version 2

The research team addressed these problems in a new version of the scale. The scoring system was modified so that the items would only refer to the use of one well-defined skill. Each item would also be scored based on skill implementation versus absence of skill. In behavioural sciences, the use of a present or absent coding format is fairly common in both pediatric35 and adult evaluations36, especially when the respondent is asked to report on another’s status. One item was removed, as we were unable to assess it from the available video recordings (how the nurses came into contact with patients was absent from our recordings). Moreover, two ambiguous items were each subdivided in two. Item descriptions and behaviour examples were further revised and simplified to ensure maximum clarity. The order of items was also rearranged to be more consistent with the timeline of the encounters. Instructions on the use of the scale were also developed as well as an additional document describing the targeted hypnotic communication techniques, so that a lay rater could use the scale without prior knowledge of hypnosis. These changes led to a second version of the scale (N = 11 items). Before proceeding to the next step, the scale was reviewed by the team and feedback as well as minor formulation edits were done.

This Version 2 was applied by the same raters used for Version 1 (JA and MPB), using the interactions previously selected to evaluate clarity, usability and understandability and remaining issues were raised. Three items needed additional specification (labelled Synchrony, Nurse’s attention and Hypnotic bubble). Raters had difficulty assigning the appropriate score, because items were still ambiguous and required too much personal interpretation to yield appropriate reliability. For example, to rate the nurses’ attention or synchrony, one rater focused more on specific behaviours while the other rater took into account a global impression of the entire encounter.

2.2.1.3. Version 3

The three items’ descriptions and examples were further adjusted in order to maximize clarity and minimize subjective interpretation as well as focus raters’ attention on observable behaviours. Following these modifications, the test version was finalized (N = 11 items). Five items dealt with the nurse-patient relationship, while six items dealt with the use of specific communication skills and techniques. We created two count subscores to reflect the number of positive items for each category and a total count score to reflect the number of hypnosis-based communication behaviours. As these variables were count scores, it was not necessary to ascertain internal consistency.37 Before completing the second
objective of the study, a final pre-test was performed using the same 10 interactions. Raters’ and the research team’s judgment on clarity, non-ambiguity and usability, was positive and we decided to proceed with further analyses.

2.2.2. Objective 2: inter-rater reliability of the final test version (Version 3)

IRR coefficients were computed in a randomly selected portion of the videos. Time points before and after training were available for this study. For this purpose, one third of the intervention study’s videos (n = 42) that had not been previously used for Objective 1 were selected and rated. A proportion of 25–50% is considered sufficient in psychological assessment. An online random number generator was used to randomly select the videos allocated across the 4-time points.

In order to study reliability, we led two studies involving the same raters as in Objective 1 (Study 1) and untrained fully independent raters (Study 2). The second study was led to check for dissemination capacity of the scale in other independent teams and with raters with a nursing training naive to hypnosis. In IRR Study 1, raters were two female psychology graduate students (ages 22 and 24). In IRR Study 2, raters were one female nurse and one male nurse (ages 50 and 47, with 15 and 25 years of experience, respectively).

2.3. Statistical analysis

All analyses were conducted using IBM SPSS Statistics 24. IRR was computed at score and subscore level as well as for each item. Intraclass correlation coefficient (ICC) analyses were conducted to quantify the level of agreement between raters for scores. For Study 1, a two-way mixed single measure (absolute agreement) ICC was used while a two-way random single measure (absolute agreement) ICC was used for Study 2. The confidence interval was set at 95%. The following guidelines were used to interpret ICC values: 0–0.40 = poor, 0.40–0.59 = fair, 0.60–0.74 = good, 0.75–1.0 = excellent. Cohen’s Kappa was used to assess each item’s chance corrected inter-rater agreement.

The kappa values were interpreted as follows: 0–0.20 = no agreement, 0.21–0.39 = minimal, 0.40–0.59 = weak, 0.60–0.79 = moderate, 0.80–0.90 = strong, above 0.90 = almost perfect agreement. Percent agreement was also computed for each item.

For informative purposes, additional analyses were performed to establish repeatability. The means and standard deviations of the differences in total score and subscores attributions for Study 1 (Rater A – Rater B) and Study 2 (Rater C – Rater D) were computed in order to determine the limits of agreement. Bland-Altman graphs as well as Kendall correlation coefficient were used to determine the magnitude of differences in score attributions. The measurement error and the error range (i.e. above and below the actual measurement) were also calculated.

3. Results

3.1. Objective 1

Following the steps detailed in the methods, a final version of the Sainte-Justine Hypnotic Communication Assessment Scale was elaborated (Appendix). The final version is composed of two categories of behaviours classified on the basis of theory. The scale includes 11 items pertaining to relationship or technical skills. The “Relationship” category consists of 5 items dealing with (1) the adjustment of the nurse’s language to the child’s age, (2) the verbal pace adopted by the nurse, (3) whether the nurse and the patient are attuned (synchrony), (4) the development of a cooperative relationship between the nurse and the patient as well as (5) the child’s position during the medical procedure. The “Technique” category refers to the use of hypnotic communication per se and is made of six items: (1) the use of the child’s different senses in the nurse’s verbal behaviours (i.e. VAKOG, or children’s visual, auditory, kinesthetic, olfactory and gustatory senses), (2) the nurse’s attention focus on the child, (3) behaviours used to support the child, (4) the use of comforting language, (5) the identification of the use of a technique taught during the training and (6) whether the child’s behaviours suggest the experience of a hypnotic bubble.

For each item, positivity is determined as the use of one skill as reflected by specific observable behaviours. Raters should make a decision on the presence (=1) or the absence (=0) of these behaviours. A “Not applicable” or don’t know score (NA) is also available if a score cannot be decided upon. As detailed in the methods, count scores are computed by adding the number of positive scores reflecting the number of favourable behaviours. Subscores are computed separately for the Relationship and Technique categories and a total score is computed from the eleven items (Appendix).

3.2. Objective 2

3.2.1. Study 1

When a randomly selected sample of 42 nurse-patient interactions were rated by psychology graduate students (raters A and B), ICCs level reflected excellent reliability for the total score (ICC = 0.924, 95% CI = 0.864-0.958) as well as for the Relationship subscore (ICC = 0.955, 95% CI = 0.916-0.975) and Technique subscore (ICC = 0.888, 95% CI = 0.802-0.938) (Table 1). When exploring reliability at the item level, we found that nine out of eleven items had either high or perfect agreement, with Kappa values ranging from 0.844 to 1.00. For two items, reliability was moderate with Kappa values of 0.656 (Pace) and 0.725 (Nurse’s attention). Although these values were lower than those for the other items, they indicated adequate agreement between raters.

Notably, percent agreements for items ranged from 93% to 100% (Table 1).

Detailed results are available for repeatability analyses in supplementary figures (Fig. S1). These indicate that a uniformity of variance in the repeated measurements was present for the total score (τ = 0.051, p = 0.691) and for both the Relationship (τ = 0.019, p = 0.691) and Technique subscore (τ = 0.051, p = 0.691).

Table 1

<table>
<thead>
<tr>
<th>Relationship items</th>
<th>ICC</th>
<th>Kappa (c)</th>
<th>Percent agreement (%)</th>
<th>Technique items</th>
<th>ICC</th>
<th>Kappa (c)</th>
<th>Percent agreement (%)</th>
<th>Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>0.876</td>
<td>97.62</td>
<td>97.62</td>
<td>Items</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pace</td>
<td>0.656</td>
<td>97.62</td>
<td>97.62</td>
<td>VAKOG’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synchrony</td>
<td>0.947</td>
<td>97.62</td>
<td>97.62</td>
<td>Nurse’s attention</td>
<td>0.725</td>
<td>95.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooperation</td>
<td>1.000</td>
<td>100</td>
<td>100</td>
<td>Support of the child</td>
<td>0.847</td>
<td>92.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child’s position</td>
<td>1.000</td>
<td>100</td>
<td>100</td>
<td>Comforting language</td>
<td>0.847</td>
<td>92.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Use of a technique</td>
<td>0.869</td>
<td>95.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relationship subscore</td>
<td>0.955</td>
<td>–</td>
<td>–</td>
<td>Hypnotic bubble</td>
<td>0.846</td>
<td>95.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Technical subscore</td>
<td>0.888</td>
<td>–</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

VAKOG’ = The use of the child’s visual, auditory, kinesthetic, olfactory and gustatory senses. ICC = Intraclass Correlation Coefficient.
p = 0.894) and the Technique subscores (τ = −0.147, p = 0.263). The error range indicated a 0.995 above or below the actual measurement for the total score as well as 0.361 and 0.887 above or below the actual measurement for the Relationship and Technique subscores, respectively.

3.2.2. Study 2

When nurses (raters C and D) rated the same sample of nurse-patient interactions, ICCs also demonstrated excellent IRR for the total score (ICC = 0.869, 95% CI = 0.769-0.927) and for both the Relationship (ICC = 0.844, 95% CI = 0.728-0.913) and the Technique subscores (ICC = 0.868, 95% CI = 0.765-0.927) (Table 2). As for the reliability at the item level, we found that nine out of the eleven items had a high or almost perfect agreement, with Kappa values ranging from 0.806 to 0.901. Inter-rater agreement was weak for two items, with Kappa values of 0.489 (Language) and 0.581 (Support of the child). Percent agreement for each item ranged from 81% to 95% (Table 2).

Detailed results are also available for repeatability analyses in supplementary figures (Fig. S2). For the total score, a significant correlation between differences and means was found (τ = 0.260, p = 0.035), suggesting that higher the scores, larger were the differences. A uniformity of variance in the repeated measurements was noted for the Relationship (τ = 0.247, p = 0.068) and the Technique subscores (τ = 0.190, p = 0.142). The error range indicates a 2.945 above or below the actual measurement for the total score as well as 1.626 and 1.581 above or below the actual measurement for the Relationship and Technique subscores, respectively.

4. Discussion

The aims of this study were to develop a scale assessing healthcare professionals’ behaviours when using HCTech and to evaluate its IRR. The SJ-HCAS was developed by a multidisciplinary team (physicians, psychologists, nurse, psychology graduate student) based on key elements of the practice identified in the literature and the nurses’ training, as well as experts’ opinions on hypnotic communication. Members were all involved in the scale development and agreed on items’ descriptions as well as examples.

To our knowledge, the SJ-HCAS is the first tool to assess pediatric healthcare professionals’ behaviours when using hypnotic communication techniques. Other measurement tools assessing nurse-patient interactions (Child-Adult Medical Procedure Interaction Scale14 and Measure of Adult and Infant Soothing and Distress45) are available in the literature. Several studies have used video-recordings for the purpose of training and assessing communication skills in oncology nurses.6,47 This study not only shows that video-based skill implementation assessments are feasible, but also that absence of skills is identifiable. This observation is consistent with the study by Birnbach et al.45 who showed that video technology helps identify inadequately learned skills and can lead to more in-depth training. The SJ-HCAS can be used as a teaching tool as it allows raters to evaluate the presence or absence of a skill. This could serve to identify teaching opportunities for hypnosis communication trainers.

Overall IRR for quantitative scores (total score and subscores), for both psychology graduate students and independent nurses, was excellent. The ICCs in Studies 1 and 2 were not significantly different (overlapping CIs). Thus, an excellent inter-rater agreement was replicated in Study 2, suggesting that the SJ-HCAS has good dissemination abilities. When looking at IRR at the item level, nine of the eleven items in both studies had a high or perfect agreement, indicating that the items’ description and examples were clear and non-ambiguous for psychology students as well as practising nurses. However, between both studies, four items had a lower agreement, suggesting that some items might be reworded or clarified for independent lay users.

In Study 1, the Pace item had a moderate agreement. This might be due to the raters’ difficulty in distinguishing between nurses who spoke slowly spontaneously and those who spoke quietly deliberately in order to comfort the patients. This created confusion when evaluating this behaviour. Regarding the Nurses’ attention, this item also had a moderate agreement. Studies have shown that nurses are frequently disturbed by different sources and types of interruptions when performing daily tasks.49,50 To score this item, raters had to take into account disruptions that occurred in the interactions (e.g. answering doctors or parents’ questions during the medical procedure) and the nurses’ reactions to these disruptions. Raters had to evaluate the nurses’ attention behaviours only based on the item’s description and examples provided. This may have created disrupting noise and may explain differences between raters.

In Study 2, we found a lower agreement for the Language item. When evaluating the nurses’ sensitivity to language, raters had to evaluate if the language was both appropriate to the child’s age, but also to his or her context of life. In pediatrics, jargon along with medical terminology can be frightening and confusing for children.51 A factor that could account for this observation is that the independent nurses who participated in this study did not have previous experience with children while the raters in Study 1 specialized in pediatrics. As a result, these nurses may not have had the knowledge to assess whether the language was adapted to the child. This could also explain why this item obtained excellent agreement in Study 1 but a weak agreement in Study 2. In regards to the Support of the child, this item also had a lower agreement. When treating patients, communication can be done with or without words by using gestures and facial expressions52: verbally supporting the child versus touching the patient. One rater may have focused on the verbal support while the other focused on the non-verbal.

Although limited, it is possible that differences in IRR between Study 1 and Study 2 may relate to differences in professional background (psychologists versus nurses). In fact, nurses had a personal experience of venipuncture which could serve as a basis for their judgment and consequently increase inter-rater discrepancies. If this is true, it seems all the more important to prompt raters to assess observed behaviours without referring to their own experience or history. It also underscores a certain degree of naivety or ingenuousness that is probably necessary to reliably rate the scale.

### Table 2

Inter-rater reliability analyses of 42 randomly selected nurse-patient interactions. (Raters: nurses).

<table>
<thead>
<tr>
<th>Items</th>
<th>ICC</th>
<th>Kappa (co)</th>
<th>Percent agreement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>0.849</td>
<td>80.95</td>
<td>1</td>
</tr>
<tr>
<td>Pace</td>
<td>0.901</td>
<td>95.24</td>
<td>1</td>
</tr>
<tr>
<td>Synchrony</td>
<td>0.901</td>
<td>95.24</td>
<td>1</td>
</tr>
<tr>
<td>Cooperation</td>
<td>0.856</td>
<td>92.86</td>
<td>1</td>
</tr>
<tr>
<td>Child’s position</td>
<td>0.893</td>
<td>95.24</td>
<td>1</td>
</tr>
<tr>
<td>Relationship subscore</td>
<td>0.844</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Technique subscore</td>
<td>0.868</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Items</th>
<th>ICC</th>
<th>Kappa (co)</th>
<th>Percent agreement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nurse’s attention</td>
<td>0.806</td>
<td>90.48</td>
<td>1</td>
</tr>
<tr>
<td>Support of the child</td>
<td>0.581</td>
<td>80.95</td>
<td>1</td>
</tr>
<tr>
<td>Comforting language</td>
<td>0.836</td>
<td>92.86</td>
<td>1</td>
</tr>
<tr>
<td>Use of a technique</td>
<td>0.897</td>
<td>95.24</td>
<td>1</td>
</tr>
<tr>
<td>Hypnotic bubble</td>
<td>0.847</td>
<td>92.86</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Items</th>
<th>ICC</th>
<th>Kappa (co)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relationship</td>
<td>0.889</td>
<td></td>
</tr>
<tr>
<td>Technique</td>
<td>0.869</td>
<td></td>
</tr>
</tbody>
</table>

*VAKOG = The use of the child’s visual, auditory, kinesthetic, olfactory and gustatory senses. ICC = Intraclass Correlation Coefficient.*
Repeatability analyses showed larger error ranges and limits of agreement when nurses assessed hypnotic communication rather than psychology graduate students. This variability may have occurred because raters from Study 1 were involved in the scale development process. Perhaps their involvement facilitated the ease of use of the scale and allowed higher measurement accuracy. Although a limited systematic bias was found for Study 2’s total score repeatability analysis, in general, the SJ-HCAS can be considered as providing repeatable results.

We should recognize limitations to this study. First, due to feasibility constraints in the outpatient pediatric oncology clinic, the sample of nurses was limited in size and selection biases may have occurred. It is possible that including a larger sample of nurses would have increased behavioural variability and thus would have provided a more realistic test of IRR. Second, although we used extant literature on developing topics and domains to target, the scale was designed to fit primarily with the content of the training that was offered in our site. The scale was developed to assess nurses’ hypnoanalgesia communication skills. Thus, other types of hypnotic communication elements are not covered by this instrument. This scale bears limited external validity and should not be used to assess alternative protocols of hypnotic communication other than the one based on hypnoanalgesia. Additionally, the use of the scale should be restricted to research as its validity remains to be studied. It is also important to mention that as various elements may participate to the process of induction, inter-ventions that stray from the protocol presented in this study do not constitute a poorer implementation of techniques. It is also possible that other rating format would be appropriate, including calling for Likert-type rating scales making it possible to express a more nuanced view on what raters will report. Finally, the scale is limited to the rating of desired behaviours. Although undesirable behaviours may have a strong impact (e.g. "well there, it won’t hurt much"), the definition and scope of these “negative” behaviours are yet to be determined. It is probable that future efficacy studies will prompt the assessment of such undesirable attitudes or behaviours. Despite these limitations, this study is the first to address the important issue of objectively evaluating hypnosis-derived communication. It also used an iterative process to warrant clear definition and limit overlapping of items and yield a reasonable IRR. Future research should address other properties including validity. If it is further supported, the scale could be used to evaluate the effect of training in this field.

5. Conclusion

We developed the first scale to rate and score hypnotic communication in nurse-patient interactions. The development followed an iterative process and yielded an 11-item scale to assess relationship quality and technique use. The results from the IRR studies support further use of the scale to evaluate hypnotic communication. The use of such an instrument bears an important impact as it could help to demonstrate that observable effects of training are associated with outcomes in professionals and patients by assessing integrity. This could promote the use of hypnosis-derived techniques in daily care. Pediatric nurses have an important role in cancer treatment and their abilities to use hypnotic communication during painful procedures has the potential to greatly diminish children’s pain and distress.

Declaration of interests

The authors declare no conflict of interests.

Acknowledgments

The authors wish to thank Marie-Pier Bilodeau for her help with the scale development process. Support for this research came from the CHU Sainte-Justine Foundation awarded to the Centre of Psycho-Oncology (Dr. Serge Sultan).

Appendix A: Supplementary data

Supplementary data associated with this article can be found in the online version, at https://doi.org/10.1016/j.ctim.2017.11.013.

References


