
Assessing the Suitability of Written Stroke Materials: An Evaluation of the Interrater Reliability of the Suitability Assessment of Materials (SAM) Checklist

Tammy Hoffmann, PhD,^{1,2} and Yvette Ladner, BOccThy (Hons)²

¹Centre for Research in Evidence-Based Practice (CREBP), Faculty of Health Sciences and Medicine, Bond University, Gold Coast, Australia; ²Division of Occupational Therapy, School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, Australia

Purpose: Written materials are frequently used to provide education to stroke patients and their carers. However, poor quality materials are a barrier to effective information provision. A quick and reliable method of evaluating material quality is needed. This study evaluated the interrater reliability of the Suitability Assessment of Materials (SAM) checklist in a sample of written stroke education materials. **Methods:** Two independent raters evaluated the materials (n = 25) using the SAM, and ratings were analyzed to reveal total percentage agreements and weighted kappa values for individual items and overall SAM rating. **Results:** The majority of the individual SAM items had high interrater reliability, with 17 of the 22 items achieving substantial, almost perfect, or perfect weighted kappa value scores. The overall SAM rating achieved a weighted kappa value of 0.60, with a percentage total agreement of 96%. **Conclusion:** Health care professionals should evaluate the content and design characteristics of written education materials before using them with patients. A tool such as the SAM checklist can be used; however, raters should exercise caution when interpreting results from items with more subjective scoring criteria. Refinements to the scoring criteria for these items are recommended. The value of the SAM is that it can be used to identify specific elements that should be modified before education materials are provided to patients. **Key words:** health education, interrater reliability, patient education, SAM, stroke, written education materials, written information

The importance and benefits of providing patients and their carers with education are widely acknowledged.¹ Written patient education materials are frequently used by health professionals and desired by patients,²⁻⁴ and they offer benefits such as portability and message consistency.⁵ Additionally, patients are able to read information at a time and speed that are suitable for them, choose how much information they receive, refer back to information as needed, and share information with family members who were not present during the consultation.⁵⁻⁷ Health care professionals are encouraged to use written education materials with patients to reinforce and supplement information that is provided orally.^{8,9} When only oral information is provided, much of that information is forgotten.¹⁰⁻¹²

Stroke patients and their carers need to receive information on a wide variety of topics and at all phases of the recovery process.¹³ Information can assist individuals to understand what has happened, cope with the consequences, know

what secondary stroke prevention behaviors to perform, and facilitate self-management of a chronic condition. Providing information to stroke patients and their carers has been found to improve patient and carer knowledge and aspects of patient satisfaction and to reduce patient depression.¹⁴ Clinical guidelines for stroke care now recommend that all patients and their families be offered information that is tailored to their needs.¹⁵ The majority of stroke patients and their carers would like to receive written information,^{16,17} although dissatisfaction with the quality of the written materials provided has been identified as a barrier to adequate stroke education by patients,¹⁸ carers,¹⁸ and stroke health professionals.¹⁹ Despite the reading difficulties that are often experienced by patients with aphasia, Rose et al²⁰ examined

the information preferences of 40 patients with aphasia and found that most wished to receive written materials.

For all patients, written health education materials can only be of benefit if they can be understood by the recipient. However, for patients with stroke, various impairments, such as aphasia and visual and perceptual difficulties, can occur after a stroke and have an impact on the effectiveness of patient education. Therefore, consideration of the content and design characteristics of the written materials that are used with stroke patients is particularly warranted.^{16,21} Written stroke materials are often presented at a reading level that is higher than that of the target audience.²¹⁻²³ However, other factors such as the design, layout, content, graphics, and overall presentation of the material play an essential role in determining whether the materials will be understood and used by those who receive them.²⁴ A study by Eames et al¹⁶ indicated that stroke patients identified factors relating to the type of language, font size, color, and graphics as important in facilitating comprehension of written stroke materials.

Few published assessments have been designed to evaluate the suitability of written health education materials.²⁵⁻²⁷ One of the most widely used assessments is the Suitability Assessment of Materials (SAM) checklist.²⁸ Despite its use, details of the interrater reliability of the individual items of the SAM have not been published. This information may assist health care professionals in making an informed decision about using the SAM to determine whether written education materials are appropriate for their target audience. The aim of this study was to evaluate the interrater reliability of the individual items and overall rating of the SAM for a sample of written stroke education materials.

Methods

Procedure

A range of written stroke education materials was gathered as part of a study¹⁷ that aimed to determine current practice in the provision of written stroke information to stroke inpatients and their carers. Participants in that study were

recruited from a large metropolitan hospital in Brisbane, Australia, and copies of the written stroke education materials that were provided to patients and carers throughout patients' hospital stays were obtained. The suitability of these written materials was independently rated with the SAM by 2 researchers (T.H. and K.M.), who were experienced in using the SAM and were experts in the suitability of written health education materials.

Suitability Assessment of Materials

The SAM consists of 22 individual items that are grouped under 6 categories: content, literacy demand, graphics, layout and typography, learning stimulation and motivation, and cultural appropriateness. Within these categories, individual items are given a rating of not applicable, 0 (not suitable), 1 (adequate), or 2 (superior) according to how well they meet the criteria for each item. The ratings for each item are then summed to yield a total SAM score. This is then converted into a percentage score by dividing the total SAM score by the total possible score for that particular material. The percentage scores are then used to classify the materials as unsuitable (0%–39%), adequate (40%–69%), or superior (70%–100%).

Written materials

A total of 25 stroke written education materials were rated. In terms of overall size, when folded out, 12 were A4 (standard business letter size: 11.7 × 8.3 inches), 8 were A3 size (twice the size of A4), and 5 were A5 size (half the size of A4). Two were double-sided A4 sheets, 10 were booklets (bound documents containing between 5 and 27 pages), 7 were brochures (single-sheet A4 documents with fold-out sections or A5 folded documents), and 6 were information sheets (folded A3 documents). When content was considered, materials were grouped as containing either general or specific stroke information. Seventeen materials provided general stroke information, such as what a stroke is and the causes and management of a stroke, while the remaining 8 contained information about specific stroke-related impairments, such as aphasia or emotional disturbances after a stroke. Seven materials were published by a national

stroke organization, 5 by the hospital, and 13 by state stroke organizations. Thirteen materials were printed in color, with the remainder in black and white. Of the 25 materials, 16 were printed on glossy paper, and of the remaining 9, 6 were printed on plain paper and 3 had glossy covers with matte inner pages.

Data analysis

Data were analyzed using STATA, version 11 (StataCorp, College Station, Texas), and interrater reliability was determined by calculating the level of agreement (percentage agreement) and weighted kappa values for each of the individual SAM items and the overall category (unsuitable, adequate, superior) that was assigned to each written material. Weighted kappa values were interpreted according to the cutoffs described by Viera and Garrett²⁹: values between 0.81 and 0.99 indicate almost perfect agreement, values between 0.61 and 0.80 indicate substantial agreement, values between 0.41 and 0.60 indicate moderate agreement, values between 0.21 and 0.40 indicate fair agreement, values between 0.01 and 0.20 indicate slight agreement, and values less than or equal to 0 indicate agreement that is less than chance.

Results

Table 1 shows the percentage total agreement and weighted kappa values for each of the 22 individual SAM items and the overall SAM category for the 25 rated materials. Of the 22 SAM items, there was 100% total agreement for 7 of the items, agreement between 95% and 99% for 8 items, between 90% and 94% agreement for 4 items, and less than 90% agreement for 3 items. For the overall SAM category, the percentage total agreement was 96%.

According to the weighted kappa values, there was almost perfect agreement (values between 0.81 and 0.99) for 7 items, substantial agreement for 3 items (values between 0.61 and 0.80), moderate agreement for 1 item (value between 0.41 and 0.60), and fair agreement for 1 item (value between 0.21 and 0.40). Three items had a weighted kappa value of 0.00, which indicates that there was nil agreement better than by chance

Table 1. Percent total agreement and weighted kappa values for the individual SAM items and the overall SAM category

SAM item and description	% Total agreement	Weighted kappa value
1. Content		
a. Purpose is evident	98	0.86
b. Content about behavior	100	1.00 ^a
c. Scope is limited	96	0.76
d. Summary or review included	100	1.00 ^a
2. Literacy demand		
a. Reading grade level	100	1.00 ^a
b. Writing style, active voice	100	1.00 ^a
c. Vocabulary uses common words	94	0.65
d. Context is given first	100	1.00 ^a
e. Learning aids via "road signs"	96	0.00 ^b
3. Graphics		
a. Cover graphic shows purpose	93	0.29
b. Type of graphics	88	0.49
c. Relevance of illustrations	90	0.79
d. List, tables, etc. explained	50	0.00 ^b
e. Captions used for graphics	98	0.83
4. Layout and typography		
a. Layout factors	98	0.88
b. Typography	92	0.84
c. Subheads (chunking) used	100	1.00 ^a
5. Learning stimulation, motivation		
a. Interaction used	99	0.95
b. Behaviors are modeled and specific	98	0.90
c. Motivation – self-efficacy	96	0.90
6. Cultural appropriateness		
a. Match in logic, language, experience	88	0.00 ^b
b. Cultural image and examples	100	1.00 ^a
Overall SAM category	96	0.60

^a Kappa value of 1.00 indicates perfect agreement between raters better than by chance.

^b Kappa value of 0.00 indicates that the observed and expected percentage agreement between raters was equal to chance.

between the observed and expected percentage agreement for the items. The remaining 7 items had a weighted kappa value of 1.00, indicating perfect agreement between raters better than by chance.²⁹ The weighted kappa value of the overall SAM category was 0.60, which indicates moderate agreement.

Discussion

This study revealed that the majority of individual SAM items had high interrater reliability, with 17 of the 22 items achieving substantial, almost

perfect, or perfect weighted kappa value scores. The overall SAM category achieved a moderate weighted kappa value of 0.60. These findings are comparable to those of Rees et al,³⁰ who analyzed prostate cancer education materials and reported weighted kappa values ranging from 0.20 to 1.00 for SAM items, and those of Kaphingst et al,³¹ who reported high interrater reliability ($r = 0.88$) of the SAM when it was used to evaluate colorectal cancer Web sites. However, because no further interrater reliability details were reported in these studies, a detailed comparison with the results of this study is not possible.

In this study, there was a trend for items that had more subjective rating criteria to have a larger variation in the scores given by the raters than items that were scored by using more objective criteria. This is in keeping with Wallace et al³² who cautioned researchers about using the SAM to evaluate the suitability of written health information materials as they believed that different raters may produce different results because of the subjective nature of the criteria for some items. Item 3a had the lowest weighted kappa value of 0.29, indicating fair agreement. This item relates to the cover graphic; to achieve a superior rating, it must meet the following criteria: is friendly, attracts attention, and portrays the purpose of the material to the audience. Item 3b (type of graphics) has scoring criteria similar to those for item 3a and had a weighted kappa value of 0.49, which indicates moderate agreement. To achieve a superior rating for this item, illustrations used must be simple, adult-appropriate line drawings/sketches and should be familiar to the viewers. Both of these items require the rater to award a score based on a subjective judgment of the relevant criteria.

Determination of whether something is friendly, attracts attention, is adult appropriate, or is familiar to viewers is highly subjective and is likely to be influenced by many factors, including the rater's sociodemographic and cultural background, clinical experience, and knowledge of the target audience. Both these items had relatively high total percentage agreements of 93% and 88%, respectively. The weighted kappa values indicate that this percentage agreement was largely due to chance. In comparison, an item such as 2a

(literacy demand) had a weighted kappa value of 1.00 and a total percentage agreement of 100%, which indicates agreement better than by chance; it is rated by using an objective rating criteria, namely, a reading grade level that is calculated with a readability formula.

Other items that had kappa scores below 0.80 include 1c (scope), 2c (vocabulary uses common words), and 3c (relevance of illustrations). As with items 3a and 3b, there is an element of the scoring criteria for each of these items that requires the rater to make a subjective decision, and therefore, the ratings may be influenced by characteristics particular to the rater. Three items (2e, 3d, and 6a) had a weighted kappa value of 0.00. The total percentage agreement for these items ranged from 50% to 96%. Weighted kappa value is a function of the difference between the observed and expected percentage agreements for an item.²⁹ It takes into account the size of each disagreement and assigns weights for these ratings where agreement cannot be observed. A kappa value of zero indicates that for these items, the observed and expected percentage agreement values were equal or any agreement observed between raters was equal to chance. Similarly, the rating for the overall SAM category had a high percentage total agreement (96%) and a moderate weighted kappa rating (0.60), indicating that a moderate portion of agreement between the raters was due to chance.

The SAM enables a numerical score (in percent) to be calculated, and this score is then used to determine the overall category (superior, adequate, or not adequate) to which the material that is being evaluated belongs. However, focusing on the total score or the category may not be the most beneficial way of using the SAM. Regardless of whether the SAM is being used to evaluate existing written material or to modify material while it is being developed, the value of this tool is that it can be used to identify which specific elements of the material are not suitable. Once this is known, these elements can be modified before the material is provided to the target audience. For example, a simple modification—such as changing font size, font type, or not using all capitals—can alter the rating for item 4b from not suitable to superior.

Paul et al²⁵ developed a checklist to assess the content and design characteristics of written materials and have shown it to be a valid tool. Forty-eight experts rated the checklist and also rated and ranked a series of pamphlets (designed to display the checklist in an applied form) along with 30 lay people. The concurrent validity of this checklist was confirmed as experts rated 97% of the items included in the list as important for developing quality written materials. This checklist includes broad categories similar to those of the SAM, but the creators have attempted to include more objective rating criteria to ensure consistency between raters. For example, despite still including a statement relating to “relevant” illustrations, the criteria for evaluating graphics also include points relating to the use of color and symbols. These criteria are easily identified by the rater as present or not present. Refining the scoring criteria for the SAM items that had weighted kappa values below 0.80 may result in improved interrater reliability of these items, and reassessment of the interrater reliability after the scoring criteria have been modified for the problematic items is important. Because psychometric properties, including interrater reliability, of Paul et al’s checklist²⁵ have not been evaluated, a comparison of its properties with the SAM would be useful.

The materials that were evaluated in this study were stroke specific, which may limit the application of this study’s results to materials that are targeted to other populations. However, even though all materials rated for this study were stroke related, they varied according to format, type of source, and target audience (ranged from targeting patients alone to providing information for family and/or carers). The content of the materials ranged from general stroke information to specific impairment information. It should be noted that both raters in this study are experts in the field of written health education materials and that the interrater reliability of the SAM items may be lower when it is used by less experienced raters. Health care professionals typically receive very little, if any, training in the preparation or use of health education materials.³³ Exploration of additional psychometric properties of the SAM would also be valuable.

The potential advantages of providing written education materials to patients and their carers are not maximized if the written materials are not suitable either in terms of their readability or factors relating to design, layout, content, and presentation. All of these factors play key roles in determining whether the materials will be used and understood by the target audience.^{24,26} It is important to assess the suitability of written materials prior to their use with patients and carers. The SAM is a quick and simple tool that can be used to do this, and this study indicates that the majority of the SAM items have high interrater reliability. However, raters should exercise caution when interpreting results from items with more subjective rating criteria (such as items 1c, 2c, 3a, 3b, and 3c). It is important that the ratings of the individual items are considered ahead of the overall category, because a material may achieve a superior rating overall, despite some key items—perhaps relating to content and literacy—being rated as unsuitable.

Whether the SAM is used during evaluation of existing resources or during the development of new ones, it can assist in the identification of specific elements that are not suitable and require modification. Modifying problematic elements of written education materials and piloting them with the target audience are important for facilitating recipients’ interest in, as well as understanding and recall of, the information that the material contains. Ensuring that stroke patients and carers are provided with quality written education materials is a goal for which all health professionals who are involved in the care of stroke patients should strive.

Acknowledgments

Funding/Support: Tammy Hoffmann is supported by a National Health and Medical Research Council of Australia (NHMRC)/Primary Health Care Research and Development (PHCRED) Career Development Fellowship with funding provided by the Australian Department of Health and Ageing.

Additional contributions: The authors thank Associate Professor Kryss McKenna, who was one of the raters of stroke materials.

REFERENCES

1. McKenna K, Tooth L. *Client Education: A Partnership Approach for Health Practitioners*. Sydney: UNSW Press; 2006.
2. McKenna K, Tooth LR, King DB, et al. Older patients request more information: a survey of use of written patient education materials in general practice. *Australas J Ageing*. 2003;22:15–19.
3. O'Mahoney PG, Rodgers H, Thomson RG, Dobson R, James OF. Satisfaction with information and advice received by stroke patients. *Clin Rehabil*. 1997;11:68–72.
4. Wiles R, Pain H, Buckland S, McLellan L. Providing appropriate information to patients and carers following a stroke. *J Adv Nurs*. 1998;28:794–801.
5. Bernier MJ. Developing and evaluating printed education materials: a prescriptive model for quality. *Orthop Nurs*. 1993;12:39–46.
6. Paul S. Hospital discharge education for patients with heart failure: What really works and what is the evidence? *Crit Care Nurse*. 2008;28:66–82.
7. McPherson CJ, Higginson IJ, Hearn J. Effective methods of giving information in cancer: a systematic literature review of randomized controlled trials. *J Public Health Med*. 2001;23:227–234.
8. Theis SL, Johnson JH. Strategies for teaching patients: a meta-analysis. *Clin Nurse Spec*. 1995;9:100–105.
9. Johnson A, Sandford J, Tyndall J. Written and verbal information versus verbal information only for patients being discharged from acute hospital settings to home [review]. *Cochrane Database Syst Rev*. 2003:1–18.
10. Weinman J. Providing written information for patients: psychological considerations. *J R Soc Med*. 1990;83:303–305.
11. Divertie V. Strategies to promote medication adherence in children with asthma. *Am J Matern Child Nurs*. 2002;27:10–18.
12. Anderson JL, Dodman S, Kopelman M, Fleming A. Patient information recall in a rheumatology clinic. *Rheumatol Rehabil*. 1979;18:18–22.
13. Eames S, Hoffmann T, Worrall L, Read S. Delivery styles and formats for different stroke information topics: patient and carer preferences. *Patient Educ Couns*. 2011;84(2):e18–e23
14. Smith J, Forster A, House A, et al. Information provision for stroke patients and their caregivers. *Cochrane Database Syst Rev*. 2008;2:CD001919. doi: 10.1002/14651858.CD001919.pub2.
15. National Stroke Foundation. *Clinical Guidelines for Stroke Management*. Melbourne: National Stroke Foundation; 2010.
16. Eames S, McKenna K, Worrall L, Read S. The suitability of written education materials for stroke survivors and their carers. *Top Stroke Rehabil*. 2003;10(3):70–83.
17. Hoffmann T, McKenna K, Worrall L, Read S. Evaluating current practice in the provision of written information to stroke patients and their carers. *Int J Ther Rehabil*. 2004;11(7):303–310.
18. Eames S, Hoffmann T, Worrall L, Read S. Stroke patients' and carers' perception of barriers to accessing stroke information. *Top Stroke Rehabil*. 2010;17(2):69–78.
19. Hoffmann T, McKenna K, Herd C, Wearing S. Written education materials for stroke patients and their carers: perspectives and practices of health professionals. *Top Stroke Rehabil*. 2007;14:88–97.
20. Rose T, Worrall L, Hickson L, Hoffmann T. Do people with aphasia want written stroke and aphasia information? A verbal survey exploring preferences for when and how to provide stroke and aphasia information. *Top Stroke Rehabil*. 2010;17:79–98.
21. Hoffmann T, McKenna K. Analysis of stroke patients' and carers' reading ability and the content and design of written materials: recommendations for improving written stroke information. *Patient Educ Couns*. 2006;60(3):286–293.
22. Estrada CA, Hryniewicz MM, Higgs VB, Collins C, Byrd JC. Anticoagulant patient information material is written at high readability levels. *Stroke*. 2000;31:2966–2970.
23. Sullivan K, O'Connor F. A readability analysis of Australian stroke information. *Top Stroke Rehabil*. 2001;7:52–60.
24. McKenna K, Scott J. Do written education materials that use content and design principles improve older people's knowledge? *Aust Occup Therap J*. 2007;54:103–112.
25. Paul CL, Redman S, Sanson-Fisher RW. The development of a checklist of content and design characteristics for printed health education materials. *Health Promot J Austr*. 1997;7:153–159.
26. Vahabi M, Ferris L. Improving written patient education materials: a review of the evidence. *Health Educ J*. 1995;54:99–106.
27. Mayneaux EJ, Murphy PW, Arnold C, et al. Improving patient education for patients with low literacy skills. *Am Fam Physician*. 1996;53:205–211.
28. Doak C, Doak L, Root J. *Teaching Patients with Low Literacy Skills*. 2nd ed. Philadelphia: JB Lippincott; 1996.
29. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005;37:360–363.
30. Rees C, Ford J, Sheard C. Patient information leaflets for prostate cancer: Which leaflets should healthcare professionals recommend? *Patient Educ Couns*. 2003;49:263–272.
31. Kaphingst K, Zanfini C, Emmons K. Accessibility of web sites containing colorectal cancer information to adults with limited literacy. *Cancer Causes Control*. 2006;17:147–151.
32. Wallace LS, Rogers ES, Turner LW, Keenum AJ, Weiss BD. Suitability of written supplemental materials available on the Internet for nonprescription medications. *Am J Health Syst Pharm*. 2006;63:71–78.
33. Hoffmann T, Worrall L. Designing effective written health education materials: considerations for health professionals. *Disabil Rehabil*. 2004;26:1166–1173.