Japanese Journal of Comprehensive Rehabilitation Science (2017)

Original Article

Development of a reaction scale for dementia patients during recreational activities: a trial study of interobserver reliability and criterion-related validity

Masaki Kamiya, OT, MS,[1] **Izumi Kondo, MD, PhD,**[1] **Aiko Osawa, MD, PhD**[1]

[1]Department of Rehabilitation, National Center for Geriatrics and Gerontology, Obu, Aichi, Japan

ABSTRACT

Kamiya M, Kondo I, Osawa A. Development of a reaction scale for dementia patients during recreational activities: a trial study of interobserver reliability and criterion-related validity. Jpn J Compr Rehabil Sci 2017; 8: 109–114.

Objective: Recreational therapy is often provided for patients with dementia, but there are few indices on the reactions of patients during therapy, making evaluation difficult. The objective of this study was to create a new observational index capable of evaluating the reactions of elderly people with dementia, and to demonstrate its utility.

Methods: The observational index developed by Ishihara et al. was modified, and its interobserver reliability was examined by two assessors who evaluated 16 patients (mean age: 80.2 ± 7.2 years, mean Mini Mental State Examination [MMSE] score: 18.3 ± 7.9) hospitalized on dementia and geriatric wards to calculate the kappa coefficients. Criterion-related validity was examined by evaluating 33 patients meeting the same conditions (mean age: 80.8 ± 8.4 years, mean MMSE score: 15.6 ± 6.8) using this index and the Observed Emotion Rating Scale to calculate Spearman's rank correlation coefficient.

Results: The examination of interobserver reliability produced kappa coefficients of 0.57 to 0.67, while the examination of criterion-related validity produced correlation coefficients of 0.56 to 0.86 ($p < 0.001$).

Conclusion: The interobserver reliability and criterion-related validity of the index examined in this study were of a sufficient level for this index to be used in clinical settings to evaluate dementia patients.

Key words: dementia, observational evaluation, recreation, reliability, validity

Introduction

According to the "Dementia Disease Medical Examination Guidelines" published by the Japanese Society of Neurology in 2010, the main nonpharmacological therapies for dementia include reality orientation (RO), reminiscence therapy, music therapy, cognitive stimulation therapy, and exercise therapy. With the exception of cognitive stimulation therapy, on the whole, none of these therapies have been proven to be effective [1]. However, a proposal was issued in 2017 to promote nonpharmacological therapies as a weak recommendation in advance of a revision to the "Dementia Disease Medical Examination Guidelines" [2]. Nonpharmacological therapy for dementia is expected to progressively develop, and the examination of indices that evaluate the outcomes of nonpharmacological therapy for dementia patients is an important task in pursuing this development.

Since 2006, short-term intensive rehabilitation for patients with dementia has been available as a new form of rehabilitation for mild dementia. The scope of this rehabilitation subsequently expanded in 2009 to include moderate and severe dementia along with changes to the system, and Toba et al. demonstrated the actual situation and outcomes of short-term intensive rehabilitation for dementia patients combining therapies that included RO, cognitive stimulation therapy, and recreational therapy [3]. Spector et al. also reported the alleviation of symptoms of depression and improvement in quality of life in cognitive stimulation therapy [4, 5]. However, these reports are unclear on whether the patients were able to actively participate in training, which makes it difficult to determine if the stimulation provided was appropriate. Elderly persons with dementia are nonetheless difficult to evaluate by questionnaire due to cognitive decline, and very few observational

indices exist to evaluate their reactions.

Lawton et al. reported the "Philadelphia Geriatric Center Affect Rating Scale," in which the duration of patients' facial expressions and reactions while participating in recreation are scored through observational evaluation [6]. A subsequent revised evaluation form was completed and currently exists as the reported "Observed Emotion Rating Scale," which is the gold standard for evaluation of reactions [7]. This scale evaluates facial expressions and reactions in five separate affects: Pleasure, Anger, Anxiety/Fear, Depression/Sadness, and Interest.

Ishihara et al. also used their own observational index in an attempt to determine if environmental differences between group homes and conventional facilities influence the cognitive and emotional functions of elderly people with dementia. This was an attempt to construct a scale measuring the changes in facial expressions, degree of participation, autonomy, and sociability, and it succeeded in confirming the positivity or negativity of subjects' interests and reactions [8]. However, the reliability and validity of this scale when used in dementia patients have not been proven.

In light of these circumstances, the objective of the present study was to modify Ishihara et al.'s index to create an observational index (hereinafter "reaction scale") capable of evaluating the reactions of elderly people with dementia, clarify the utility of the scale itself, and determine the utility of recreation using the scale.

Overview of the National Center for Geriatrics and Gerontology (NCGG) Reaction Scale

Creation of the NCGG Reaction Scale

The observational scale used by Ishihara et al. in their earlier study evaluates the changes in facial expressions, degree of participation, autonomy, and sociability on a four-point scale [8]. As a pilot study for the present study, two assessors evaluated dementia patients using this scale to examine concordance. The results showed that the concordance of changes in facial expressions and sociability was lower than that of the degree of participation and autonomy. Furthermore, their examination of the number of concordant points also showed a greater number of discrepancies in the changes in facial expressions and sociability. Apathy is generally known to be the most common behavioral and psychological symptom of dementia in Alzheimer's disease [9] and is characterized by flat emotions (blunted affect) and a decline in sociability [10]. The reasons for the large number of discrepancies could be the variation in the severity of dementia in hospitals and the variation in patients who present with apathy versus those who do not.

However, it is unlikely that the severity and symptoms of dementia are consistent across hospitals and facilities,

and closer examination of the results of this pilot study suggests that a more fitting evaluation of dementia patients should be possible by measuring just the degree of participation and autonomy, which had a higher concordance and greater number of concordant points. "Encouragement" is also often necessary in elderly people with dementia due to their difficulty understanding or hearing the content of tasks as a result of cognitive decline. Thus, the decision was made to always provide controlled "encouragement" when an elderly person with dementia showed no reaction to intervention. Additional notes were included so that the assessors understood what this "encouragement" should specifically involve.

In this manner, a modified reaction scale capable of more accurately evaluating the reactions of dementia patients was created from Ishihara et al.'s index (Figure 1).

Overview of Recreation at Our Hospital

At our hospital, inpatients on the dementia and geriatric wards undergo recreational therapy with a therapist for 40 min once a week. The patients gather as a group in a wide space such as a hall to engage in recreation with the support of 8−10 occupational therapists, physical therapists, speech-language-hearing therapists, and nurses. The contents are activities that incorporate RO, exercise therapy, and cognitive stimulation therapy. Due to the varying severity of dementia among the patients, the content of recreation is not always tailored to each individual patient, but the patients participate in activities at their own pace. Four types of activity were set for the present evaluation: a word recall task, radio calisthenics, balloon volleyball, and bowling. These four activities were evaluated using the reaction scale.

Subjects and Methods

The reaction scale was examined for its reliability and validity.

1. Examination of interobserver reliability
1.1 Subjects
The subjects were 16 patients (10 men and 6 women) capable of participating in recreation on the dementia and geriatric wards of our hospital. The patients' ages ranged from 64 to 90 years (mean: 80.2 ± 7.2 years), and the Mini Mental State Examination (MMSE) scores for cognitive function ranged from 7 to 30 (mean: 18.3 ± 7.9).

1.2 Methods
In addition to the author, two experienced occupational therapists (11 and 14 years of experience) acted as assessors. During recreation, the two assessors used the reaction scale to simultaneously evaluate the same patients on an individual basis. The results of the two

Figure 1. The reaction scale used in this study.

assessors' evaluations were used to examine the concordance of each activity and the recreation overall.

1.3 Statistical analysis

SPSS Ver. 21 was used for statistical analysis. Interobserver reliability was examined using kappa coefficients, which show the rate of concordance.

2. Examination of criterion-related validity
2.1 Subjects

The subjects were 33 patients (14 men and 19 women) capable of participating in recreation on the dementia and geriatric wards of our hospital. The patients' ages ranged from 59 to 96 years (mean: 80.8 ± 8.4 years), and the MMSE scores for cognitive function ranged from 1 to 27 (mean: 15.6 ± 6.8).

2.2 Methods

In addition to the author, two experienced occupational therapists (5 and 9 years of experience) acted as assessors. During recreation, one assessor evaluated the patients using the recreation scale, while the second assessor evaluated the same patients using the Interest affect of the Observed Emotion Rating Scale. The patients were evaluated individually at the same time. The Interest affect of the Observed Emotion Rating Scale was used to show a correlation with the reaction scale and was also used as an index to prove validity. After evaluation, the two assessors' results were examined.

2.3 Statistical analysis

SPSS Ver. 21 was used for statistical analysis. To examine criterion-related validity, Spearman's rank correlation coefficient was used to calculate the correlation coefficient between the reaction scale and the Interest affect of the Observed Emotion Rating Scale.

Results

The clinical characteristics of the patients by examination item are shown in Table 1.

1. Interobserver reliability

In Table 2, the concordances between the two assessors for the degree of participation and autonomy are shown for each activity (word recall task, radio calisthenics, balloon volleyball, bowling, and recreation overall). The kappa coefficients for each activity and overall were 0.57, 0.65, 0.57, 0.66, and 0.67, respectively.

2. Criterion-related validity

Table 3 shows the correlations between the reaction scale and Observed Emotion Rating Scale. Significant correlations were seen in all activities during recreation. The correlations between the Observed Emotion Rating Scale and the word recall task were $r = 0.60$ for degree of participation and $r = 0.56$ for

Table 1. Clinical characteristics of patients by examination item.

	Examination of interobserver reliability	Examination of criterion-related validity
Sample size (n)	16	33
Men/women (n)	10/6	14/19
Age, range (years)	64−90	59−96
Age, mean ± SD (years)	80.2 ± 7.2	80.8 ± 8.4
MMSE total score, range	7−30	1−27
Mean ± SD (points)	18.3 ± 7.9	15.6 ± 6.8

MMSE, Mini Mental State Examination.

Table 2. Interobserver evaluations of each activity using the reaction scale.

Word recall task (kappa = 0.57)

2nd assessor	1st assessor			
	1	2	3	4
1	6	2	1	
2		1	2	
3			3	
4			1	6

Radio calisthenics (kappa = 0.65)

2nd assessor	1st assessor			
	1	2	3	4
1	6	1	1	
2		2	1	
3			1	2
4				12

Balloon volleyball (kappa = 0.57)

2nd assessor	1st assessor			
	1	2	3	4
1	6			
2				
3		1	1	4
4				14

Bowling (kappa = 0.66)

2nd assessor	1st assessor			
	1	2	3	4
1	6			
2		2		
3				
4			4	14

Recreation overall (kappa = 0.67)

2nd assessor	1st assessor			
	1	2	3	4
1	24	3	2	
2		5	3	
3		1	5	6
4			5	46

Table 3. Correlations between the reaction scale and the Observed Emotion Rating Scale.

	Word recall task		Radio calisthenics		Balloon volleyball		Bowling	
	Degree of participation	Autonomy	Degree of participation	Autonomy	Degree of participation	Autonomy	Degree of participation	Autonomy
Observed Emotion Rating Scale	0.60*	0.56*	0.87*	0.80*	0.86*	0.83*	0.66*	0.60*

Spearman's rank correlation coefficient (*$p < 0.001$).

autonomy ($p < 0.001$). The correlations between the Observed Emotion Rating Scale and radio calisthenics were $r = 0.87$ for degree of participation and $r = 0.80$ for autonomy ($p < 0.001$). The correlations between the Observed Emotion Rating Scale and balloon volleyball were $r = 0.86$ for degree of participation and $r = 0.83$ for autonomy ($p < 0.001$). The correlations between the Observed Emotion Rating Scale and

bowling were $r = 0.66$ for degree of participation and $r = 0.60$ for autonomy ($p < 0.001$).

When the items in the reaction scale were compared, the correlation coefficients for degree of participation were slightly higher than those for autonomy. When focus was placed on the content of the activity, the two scales were judged to have a "fairly strong correlation" in "radio calisthenics" and "balloon volleyball." These activities were performed by all patients as a group, and, although they involve dynamic exercise, the correlation coefficients were high. The reaction scale and Observed Emotion Rating Scale were judged to have a "fair correlation" in the "word recall task" and "bowling."

Discussion

1. Interobserver reliability

Unlike other diseases, the symptoms of dementia vary greatly, and dementia patients are prone to developing conditions such as circadian rhythm disorder [11]. An observational index capable of responding to the variations in symptoms is therefore needed to properly evaluate dementia patients. However, no index currently exists that is both useful and versatile. If the observational index developed and used by Ishihara et al. [8] in an earlier study could be modified to meet the traits of dementia patients and be confirmed as highly reliable, accurate evaluation of the reactions of dementia patients should become feasible. The NCGG Reaction Scale was consequently developed.

The present study demonstrates that this scale has fair or moderate to good reliability, given that the kappa coefficients calculated for each activity meet the general criteria for kappa coefficients (< 0.4 = poor, $0.4–0.6$ = fair or moderate, $0.6–0.8$ = good, >0.8 = excellent). In effect, this study demonstrated the strength of the interobserver reliability of the NCGG Reaction Scale. Activities that are commonly performed in recreation were also examined and found to have a high concordance, especially "radio calisthenics," "bowling," and "recreation overall." One of the characteristics of "radio calisthenics" is its great familiarity to this elderly generation. This concordance between the assessors' evaluations of radio calisthenics may have been high because this activity is performed simultaneously by everyone in a group. Meanwhile, "bowling" is characterized by a clear delineation between success and failure based on the number of pins that are toppled; patients' reactions to this are also clear and might explain the ease with which the assessors could make accurate observations. Together, these activities likely resulted in a high concordance for "recreation overall."

However, after performing this study, several points of caution were identified as necessary to improve the accuracy of concordance in evaluations, i.e. to perform more accurate evaluations. Specifically, in evaluations designed to understand the reactions of patients, assessors must evaluate the facial expressions and reactions of patients being evaluated each time an activity is performed. This is because the results of personal preferences or the physical condition of the dementia patient being evaluated vary between each activity. Thus, to more accurately understand the reactions of dementia patients, it is important to segment the reactions for evaluation in each activity instead of attempting to evaluate recreation as a whole.

2. Criterion-related validity

The NCGG Reaction Scale developed in this study needs to be proven as valid in order to confirm if it is an accurate evaluation index. The Face Scale [12] already exists as a metric to observe the positivity or negativity of a patient's facial expressions; however, this scale cannot evaluate movement. Furthermore, the Mini Emotional State Examination (MESE) [13] measures emotional functions including human empathy, morality, social common sense, and comfort and discomfort associated with happiness or unhappiness, based on evaluations of reactions within the five senses and reactions to questions. However, the MESE cannot evaluate reactions to recreation or the presence of autonomy. That is why the Observed Emotion Rating Scale, which is the gold standard for observational evaluation in the United States, was used as an index of validity in the present study.

The Observed Emotion Rating Scale has been proven to have adequate reliability and validity [7]. The scale was developed to evaluate an elderly person's facial expressions to determine if special care was resulting in a positive or negative affect [6]. The Observed Emotion Rating Scale is similar to the reaction scale developed in the present study in that it is an observational evaluation to determine the positivity or negativity of a patient's reactions. However, the NCGG Reaction Scale differs from the Observed Emotion Rating Scale in terms of its development history and objectives in a few areas: it evaluates a patient's physical reactions, communication within a group, and autonomy in activities, as well as facial expressions. It can therefore be considered a new index. The Observed Emotion Rating Scale has been used in the United States in the evaluation of the first therapeutic game and in Europe in the evaluation of music therapy [14, 15]. Despite being the gold standard, the Observed Emotion Rating Scale has not yet been translated into Japanese and is consequently not widely used in Japan. In light of this, the NCGG Reaction Scale, modified from Ishihara et al.'s index, is expected to be widely used in Japan as a similar observational index to the Observed Emotion Rating Scale that is capable of evaluating the positivity or negativity of reactions.

3. Study limitations and future plans

The sample size in this study was small because the evaluations were done during multiple recreation sessions at our hospital. Furthermore, because the patients were hospitalized on wards at our hospital, their dementia was moderate, with MMSE scores ranging from 15 to 18. Another study is therefore needed with a larger sample size and a wider range of dementia severity. Plans are also underway to examine the utility of each recreational activity at facilities other than our hospital and to continue to examine the content validity of this scale.

References

1. Japanese Society of Neurology. Dementia Disease Medical Examination Guidelines 2010 Compact Version 2012. Tokyo: Igaku-Syoin; 2012. p. 67−8.
2. Japanese Psychogeriatric Society. Dementia Disease Medical Examination Guidelines 2017 (plan). Available from: https://www.jpn-geriat-soc.or.jp/info/topics/pdf/2016 1116_01_01.pdf (cited 2017 May 20).
3. Japan Association of Geriatric Health Services Facilities. Verification and research on the practice and effect of dementia short-term intensive rehabilitation. Geriatric health business promotion expenses subsidy report 2007. 2008.
4. Spector A, Orrell M, Davies S, Woods RT. Can reality orientation be rehabilitation? Development and piloting of an evidence-based programme of cognition-based therapies for people with dementia. Neuropsychol Rehabil 2001; 11: 377−9.
5. Spector A, Thorgrimsen L, Woods B, Royan L, Davies S, Butterworth M, et al. Efficacy of an evidence-based cognitive stimulation therapy programme for people with dementia: crandamised controlled trial. Br J Psychiatry 2003; 183: 248−54.
6. Lawton MP, Van Haitsma K, Klapper JA. Observed affect in nursing home residents with Alzheimer's disease. J Gerontol B Psychol Sci Soc Sci 1996; 51: 3−14.
7. Lawton MP, Van Haitsma K, Perkinson M, Ruckdeschel K. Observed affect and quality of life in dementia: Further affirmations and problems. J Ment Health Aging 1999; 5: 69−81.
8. Ishihara I, Numoto K. Study of environmental factors for care of demented elderly men. J UOEH 2001; 23: 191−201.
9. Shimabukuro J, Awata S, Matsuoka H. Behavioral and psychological symptoms of dementia characteristic of mild Alzheimer patients. Psychiatry Clin Neurosci 2005; 59: 327−36.
10. Boyle PA, Malloy PF. Treating apathy in Alzheimer's disease. Dement Geriatr Cogn Disord 2004; 17: 91−9.
11. Cohen GD, Firth K, Biddle S, Lewis ML, Simmens S. The first therapeutic game specifically designed and evaluated for Alzheimer's disease. Am J Alzheimer's Dis Other Dement 2009; 23: 540−51.
12. Huskisson EC. Measurement of pain. Lancet 1974; ii: 1127−31.
13. Fujii M, Butler JP, Hirazakura A, Sasaki H. Mini-Emotional State Examination for dementia patients. Geriatr Gerontol Int 2014; 14: 508-13.
14. Hammar LM, Emami A, Gotell E, Engstrom G. The impact of caregivers' signing on expressions of emotion and resistance during morning care situations in persons with dementia: an intervention in dementia care. J Clin Nurs 2011; 20: 969−78.
15. Nakamura K, Meguro K. Relation between dementia and circadian rhythm disturbance. Nippon Rinsho 2014; 72: 579−84.