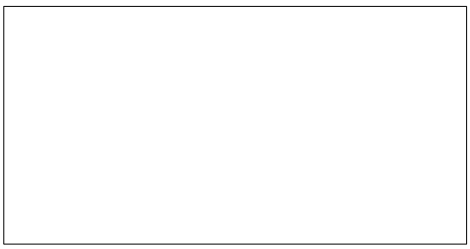


The Structured Observational Test of Function (SOTOF)

(2nd edition)

Test Manual

Alison J. Laver-Fawcett

& Eden Marrison

To cite this publication: Laver-Fawcett, A.J., and Marrison, E. (2016) The Structured Observation Test of Function (SOTOF) 2nd edition: Test Manual. York: York St John University.

Contents

List of Figures and Tables

Preface 7

Dedication 9

About the Authors 9

Acknowledgements 10

1. Introduction **11**
Summary
Dynamic assessment
Critical appraisal of the SOTOF (1st edition)

2. Administration and Scoring **16**
General introduction to the content and administration of the SOTOF Test environment, materials and equipment
Using the SOTOF Instruction Cards
Task and instruction
Possible area of deficit
Applying the SOTOF Graduated mediation protocol: Dynamic assessment
Further prompts/assessment
The Screening Assessment
Scoring the SOTOF
The four ADL tasks
Description of ADL task test items
Formats of instructions that can be selected for specific client groups
Administering the ADL tasks
Scoring the ADL tasks
The Neuropsychological Checklist and Summary Form
Completing the Neuropsychological Checklist
Normative Standards
Time
Descriptive responses
Guidelines for observing and scoring the neuropsychological test items

3. Development, Standardisation and Technical Data **41**
Content development
Rationale for the development of the SOTOF
Rationale for an observational testing format
Development of the ADL domains covered by SOTOF
Development of the constructs addressed by SOTOF
The behaviours that represent the constructs evaluated by SOTOF
Test specifications
Standardisation
Validity, reliability and clinical utility studies

Content, construct, criterion-related and face validity
Internal consistency
Test-retest and interrater reliability
Clinical utility

4. **The Test-retest and Inter-rater Reliability of SOTOF (1st edition)** **59**
Summary
Types of reliability
Methods for evaluating reliability
Reliability study: Research Questions
Methodology
First Reliability Study: Identifying and sampling the population
Procedure
Second Reliability Study: Identifying and sampling the population
Procedure
Description of sample and testing situation for the first study
Description of sample and testing situation for the second study
Description of Statistical analysis
Results
Comparison of the values obtained by each of the statistical analyses
Conclusion

5 Examination of the validity and clinical utility of SOTOF (1st edition) **75**
Defining the Concept of Clinical Utility
Procedure
The development of the questionnaires
Results
Description of samples
Results of the occupational therapist survey: Content validity
Clinical Utility

6 Enhancing the Dynamic Assessment Phase of the SOTOF (2nd edition) **94**
Aims and objectives
Methodology
Data analysis
Results
Dynamic assessment applied to neurological assessment
Dynamic assessment in relation to cognition
Review of four dynamic assessment tools
How occupational therapists use dynamic assessment
Discussion
Application of findings to the SOTOF
Relevance for occupational therapy practice and research
Strengths and Limitations
Conclusion
Recommended future research

Appendix **106**

Normative Standards Tables 106

Enlarged Written Instructions 111

Glossary 129

Further Reading and Other Assessment 133

References 135

List of Figures and Tables

Figures

Figure 1.1: The SOTOF assessment model - the relationship between occupational performance and neurological deficits

Figure 2.1: Start of the Screening Assessment Instruction Card (showing the four types of information columns).

Figure 2.2: Example of first part of the Screening Assessment Record Form

Figure 2.3: First part of the Neuropsychological Checklist

Figure 5.1: Graph showing Overall Time Taken by Subjects to complete SOTOF

Figure 5.2: Graph showing comparison of times taken across SOTOF sub-tests

Figure5.3: Graph showing the Time Taken on the Screening Test

Figure 5.4: Graph showing the Time Taken by subjects to complete the Eating Task

Figure 5.5: Graph showing the Time Taken by subjects to complete the Washing Task

Figure 5.6: Graph Showing Time Taken by subjects to complete the Drinking Task

Figure 5.7: Graph showing the Time Taken by subjects to complete the Dressing Task

Tables

Table 2.1: Equipment and materials required for SOTOF

Table 2.2: Six-level Graduated Mediation Protocol

Table 2.3: Summary of research studies to evaluate the psychometric properties and clinical utility of SOTOF

Table 2.4: Summary scores for the four ADL Tasks

Table 2.5: Overall Score for each of the four ADL tasks

Table 3.1: Summary of research studies to evaluate the psychometric properties and clinical utility of SOTOF

Table 3.2: Summary of research studies to evaluate the psychometric properties and clinical utility of SOTOF

Table 3.3: Key to concurrent and predictive measures

Table 3.4: Summary of research studies to evaluate the psychometric properties and clinical utility of SOTOF

Table 3.5: Summary of research studies to evaluate the psychometric properties and clinical utility of SOTOF

Table 4.1: The Average percent agreement for test-retest reliability for the SOTOF

Table 4.2: The Average percent agreement for inter-rater reliability for the SOTOF

Table 4.3: Average percent agreement for test-retest reliability for the SOTOF Neuropsychological Checklist

Table 4.4: Average percent agreement for inter-rater reliability for the SOTOF Neuropsychological Checklist

Table 4.5: Significance of inter-rater and test-reliability for the SOTOF Screening Assessment and Four ADL Tasks

Table 4.6: Significance of Reliability of Neuropsychological Checklist Items

Table 4.7: approximate average Kappa values for the Screening Test and four ADL Tasks for test-retest reliability

Table 4.8: approximate average Kappa values for the Screening Test and four ADL Tasks for inter-rater reliability

Table 4.9: Approximate average Kappa values for the Neuropsychological checklist for test-retest reliability

Table 4.10: Approximate average Kappa values for the Neuropsychological checklist for inter-rater reliability

Table 5.1: Therapists' Grades (n = 44)

Table 5.2: Therapists Current Clinical Area (n = 44)

Table 5.3 Content Validity - Therapists perception of the content of SOTOF (n = 44). A Comparison of results to questions 3 and 18

Table 5.4: Therapists' perceptions of the SOTOF instructions, protocols and forms (n=44)

Table 5.5: Therapists' perceptions of SOTOF materials (n = 44)

Table 6.1: Search strategy 1

Table 6.2: Search strategy 2

Table 6.3: Search strategy 3

Table 6.4: Justification for databases chosen

Table 6.5: Types of literature included and reviewed

Tables in the Appendix

Table A.1: Performance of normal sample - time taken on sub-tests
Table A.2: Time taken by normal sample on Screening Assessment (n=64)
Table A.3: Time taken by normal sample on Task 1 (n=67): eating from a bowl using a spoon
Table A.4: Time taken by normal sample on Task 2 (n=68): washing hands in a bowl
Table A.5: Time taken by normal sample on Task 3 (n=70): pouring and drinking
Table A.6: Time taken by normal sample on Task 4 (n=64): putting on a shirt
Table A.7: Total time taken by normal sample on SOTOF (n=41): Screening Assessment and four tasks
Table A.8: Descriptions of use of objects given by normal elderly clients for Task 1
Table A.9: Descriptions of use of objects given by normal elderly clients for Task 2
Table A.10: Descriptions of use of objects given by normal elderly clients for Task 3
Table A.11: Descriptions of use of objects given by normal elderly clients for Task 4
Table A.12: Descriptions of taste of food given by normal elderly clients for Task 1
Table A.13: Descriptions of taste of drink given by normal elderly clients for Task 3
Table A.14: Descriptions of temperature of water given by normal elderly clients for Task 2

Note: Throughout this manual, the person administering the test is referred to as the 'tester' , ‘rater’, ‘assessor’ or 'therapist', and the person who is being assessed as ‘patient’, 'client', ‘participant’ or ‘person’.

Preface

The **Structured Observational Test of Function** (SOTOF; Laver & Powell, 1995) is a reliable, quick and portable assessment of self-care and neuropsychological function in older people, for use in the client's home and in clinical settings (e.g. in-patient wards, out-patient clinics, occupational therapy departments, day hospitals, day centres and nursing homes). The 1st edition of SOTOF was developed over a four-year period in the 1990s. SOTOF simultaneously evaluates an older person's performance of four crucial activities of daily living (ADL; eating, washing, drinking and dressing) and generates information related to underlying perceptual, cognitive, sensory and motor performance components. It can be used to evaluate the impact of neurological deficits (e.g., resulting from cerebrovascular accidents, dementia, head injury and Parkinson's disease) on older people's ability to perform personal ADL. It comprises standardised and dynamic assessment phases. SOTOF offers a structure to observe and evaluate performance of four activities of daily living and provides a profile of the person's related skills and deficits. It can be used to signpost the need for further assessment and help inform goal setting and intervention planning. Occupational therapists, clinical psychologists and neuropsychologists should find it useful when assessing people's skills and deficits and developing plans for intervention and care management.

The development of the four SOTOF ADL tasks was based on detailed activity analysis, a survey of occupational therapy assessment practice, literature review and critique of other assessments (Laver, 1994). Each ADL task is broken down into test items which represent discrete behaviours which are scored using a dichotomous 'able' or 'unable' rating. Standardised instructions are provided for all test items. If the person is unable to successfully complete an item then the therapist is guided to use a diagnostic reasoning process to form hypotheses regarding the underlying cause of dysfunction. Suggested prompts, cues and further assessments are outlined for each test item in the SOTOF instruction cards. The most common neuropsychological deficits associated with the failure to perform a test item are provided on the SOTOF Instruction cards to aid initial hypothesis generation (Laver & Powell, 1995). In addition to recording whether a person has been able or unable to successfully complete the test item, the therapist also records qualitative information related to performance, hypothesised deficits and the person's response to any prompts or cues used.

Studies established construct, concurrent and criterion-related validity, face validity, clinical utility and internal consistency, and indicated acceptable levels of test-retest and inter-rater reliability for the SOTOF screening assessment, the four activities of daily living (ADL) tasks and the neuropsychological checklist (Laver, 1994; Laver & Powell, 1995). SOTOF can be used as both a criterion-referenced and a norm-referenced test. Normative standards, for the time taken to undertake the standardised phase of the test and for responses to items requiring the person to provide a verbal description, were established. The SOTOF was found to discriminate between patients with neurological impairment and healthy older adults (Laver, 1994; Laver & Powell, 1995).

Dynamic assessments provide analysis of abilities and how these can change owing to mediation (prompts, cues, modification, demonstration, assistance). In this 2nd edition, the dynamic phase of the assessment has been formalised and enhanced through the addition of a six-level graduated mediation protocol. If a person is unable to successfully do a SOTOF test item, then the assessor follows the graduated mediation protocol to identify the level of mediation that best facilitates the person's function in this test area. In the 2nd addition, the SOTOF instruction cards have been enhanced with detailed examples for four of the mediation levels for each test item. In addition, a six-point ordinal scoring system has been added and the SOTOF scoring forms updated. A literature review has also been undertaken, by Kathryn Wall, to update the Glossary in this manual.

This Test Manual provides a full introduction to the SOTOF; together with all the information that you will need for test administration and scoring the test. As some older people may experience problems with hearing and some people benefit from additional written instructions, a set of enlarged written instructions are provided towards the end of this manual. These instructions may be photocopied / printed as required. If you work with people with dementia or with hearing deficit frequently, then we advise that you print and laminate a set of the enlarged instruction cards. The manual also describes the development and standardisation of the SOTOF and summarises results from a number of psychometric studies conducted to evaluate the reliability, validity and clinical utility of the test.

As the SOTOF involves activities using food, water and a drink you may find it beneficial to laminate the Instruction Cards in case they get splashed during the test administration.

We are very interested to hear from anyone who is using, or is interested in using, the SOTOF for research purposes. Feedback from people using the SOTOF in clinical practice is also extremely welcome.

All correspondence regarding the SOTOF should be sent to

Eden Marrison at: **edenmarrison@yahoo.co.uk**

Alison Laver-Fawcett has now retired as of 9.10.2025 you can contact her via ResearchGate: https://www.researchgate.net/profile/Alison_Laver_Fawcett

Dedication

I wish to dedicate this assessment with grateful thanks to: my parents, Diane and Philip Laver and my children Lucas and Beatrix; and to the many individuals who participated in the studies and whose generous time commitment enabled the development of the SOTOF.

About the authors

Alison Laver-Fawcett, PhD, OT(C), DipCOT, PCAP, SFHEA

Alison has worked as a Senior Lecturer at York St John University since September 2008. She qualified from Dorset House School of Occupational Therapy, Oxford, as an occupational therapist in 1986 and completed her PhD in Psychology from the University of Surrey, Guildford, in 1995. She developed the Structured Observational Test of Function during her doctoral studies (1990- 1995). She has worked as a clinician, researcher, educator and professional lead in the UK. She undertook a one year post-doctoral fellowship at Washington School of Medicine, St Louis, USA (1993-1994). Alison then worked as an Assistant Professor at McMaster University, Hamilton, Canada, for 3.5 years (1994-1997). Following a year and a half at the University of Teesside (1998-1999), Alison moved back into the NHS as a Mental Health Project Worker (2000-2003), and was then appointed as a modernisation manager leading service evaluation and improvement for Older Peoples' Mental Health Services (2003-2008). She has undertaken consultation projects for the Care Services Improvement Partnership, Royal British Legion and the National Patient Safety Agency. Her current research relates to the development of the Activity Card Sort – UK version and an NHIR funded study related to: 'Models of Reablement Evaluation: a mixed models evaluation of a complex intervention'. She has recently led service evaluations for the York 'Ways to Wellbeing' Social Prescribing Service and for Healthwatch York. She has published a number of articles and book chapters and is author of 'Principles of Assessment and Outcome Measurement for Occupational Therapists and Physiotherapists: Theory, Skills and Application' (2007). She has previously served as the Council Member for International Affairs for the British Association of Occupational Therapists and College of Occupational Therapist; in this role she represented BAOT/COT as the UK delegate on the World Federation of Occupational Therapists Council (2009-2011). She has been a Senior Fellow of the Higher Education Academy since April 2016.

Eden Marrison, BHSc(Hons)

Eden graduated from the Occupational Therapy programme at York St John University in the summer 2016. She worked on the development of the 2nd edition of the SOTOF for one year for a Student as Co-Researcher (SCoRe) project. Eden is now working as a Band 5 Occupational Therapist, York Teaching Hospital NHS Foundation Trust, York.

Acknowledgements

The Structured Observational Test of Function was conceived, initiated and developed over a five-year period (1990-1995) by Alison Laver whilst working as an occupational therapy clinician, at Bolingbroke Hospital, researcher at St George's Hospital Medical School and researcher / lecturer Canterbury Christ Church College. The work was undertaken by Alison as part of a doctoral programme of study at the Department of Psychology, University of Surrey, under the supervision of Dr Graham Powell. Many thanks are given to: Professor Diane Cox for her ongoing support, inspiration and friendship; Julia Gosden for sharing her occupational therapy practice and for clinical supervision from April 1990 to September 1991; and Professor Millard for providing research supervision to Alison whilst she was based in the Geriatric Department at St George's Hospital Medical School, Tooting, London, from April 1990 to September 1991. This work is indebted to the generous organisations who awarded the following grants: South-West Thames Regional Health Authority Locally Organized Research Grant (April 1990 - 1991); St George's Hospital Special Trustees Research Grant (April to September 1991); The Nuffield Foundation Major Research Grant (January 1992 - January 1993); The Kings Fund (Educational Bursaries in 1990 and 1991); The College of Occupational Therapists (Pressalit Award in 1991); The Constance Owen Fund (January 1993); and Canterbury Christ Church College (allocation of PCFC funding in 1992). Very many thanks to all the older people who volunteered as research participants during the development of SOTOF. The standardisation of SOTOF could not have been completed without the contribution of the occupational therapy research assistants, who conducted the SOTOF numerous times for some of the psychometric studies: Annabel Burn, Sue Angus, Tracey Bishop and Barbara Neville and to the occupational therapists who piloted the SOTOF in their clinical practice and evaluated its clinical utility through a survey. Grateful thanks to the following colleagues for their support and encouragement during the development of the 1st edition: Monica Perlmutter, for sharing her work on the *IWJ Functional Assessment* and giving permission to adapt an aspect of its scoring system for use in the 1st edition of the SOTOF Neuropsychological Checklist; Dr Carolyn Baum for her mentorship and encouragement, for editorial review of aspects of the original test manual and permission to use our assessment model in this text. Dr Jan Duchek for editorial review of aspects of the original test manual. Thank you to colleagues from Bolingbroke Hospital, St George's Hospital, Nunnery Fields Hospital, Putney Hospital, Wolfson Medical Rehabilitation Centre and Queen Mary's Hospital who provided advice and support, in particular thanks are given to: Dr Martin Bland for statistical advice; Dr Martin Brown for advice on neurological disorders; Liz Thompson, Lydia Coltman, Sara Dunn, Jill Cooper, Caroline Douglas-Cooper, Karen Douglas and Shirley Homan for undertaking the activity analyses; Dr Marion Hildick-Smith, Dr Andrew Heller and their colleagues in Canterbury for assisting with an ethical approval application and supporting the recruitment of participants for the reliability studies. The following colleagues also provided support along the way Beryl Steeden, Dr Michael Whitelaw, Jenny Powell, Sandra Huchison, Alison MacGregor, Daphne Piegrome, Paul Higgs and Dr Robert Almli. The authors for this 2nd edition would like to thank Stephen Wey, for his ideas and feedback on the additional parts developed for enhancing the dynamic assessment phase in the 2nd edition of the SOTOF, and Kathryn Wall for her work updating the glossary.

1 Introduction

Summary: The Structured Observational Test of Function (SOTOF) is a standardised assessment which evaluates the impact of neurological deficits (e.g., resulting from cerebrovascular accidents, dementia, head injury and Parkinson's disease) on older people's ability to perform personal activities of daily living (ADL). SOTOF offers a structure to observe and evaluate performance of four activities of daily living (eating, washing, pouring and drinking and dressing) and generates information related to underlying perceptual, cognitive, sensory and motor performance components. The SOTOF provides a profile of the person's skills and deficits. It can be used to signpost the need for further assessment and help inform goal setting and intervention planning. The development of the four SOTOF ADL tasks was based on detailed activity analysis, occupational therapy assessment practice, literature review and critique of other assessments (Laver, 1994). Each ADL task is broken down into test items which represent discrete behaviours which are scored using a dichotomous 'able' or 'unable' rating. The tester indicates on the SOTOF record form whether the client is able to perform the task independently, what skills (for example reaching, sequencing) are intact, what problems are evident (for example perceptual, motor) and what the underlying dysfunction (for example ideomotor apraxia, agnosia) might be. Standardised instructions are provided for all test items. If the person is unable to successfully complete an item then the therapist is guided to implement a dynamic assessment using a diagnostic reasoning process to form hypotheses regarding the underlying cause of dysfunction. Suggested mediation (prompts, cues, physical assistance and modifications to the task or environment) and potential further assessments are outlined for items in the SOTOF instruction cards. The most common neuropsychological deficits associated with the failure to perform a test item are provided on the SOTOF Instruction cards to aid initial hypothesis generation. In addition to recording whether a person has been able or unable to successfully complete the test item, the therapist also records qualitative information related to performance, hypothesised deficits and the person's response to the graduated mediation protocol which is applied to any test items they are unable to perform independently.

The SOTOF is founded on an interdisciplinary conceptual framework drawn from neuropsychological, occupational therapy and general systems theory. This conceptual framework has been applied to the tasks of eating, washing, drinking and dressing using activity analysis to extrapolate the skills, performance and neuropsychological components of these tasks as the basis of a test of function. The components of the SOTOF include Instruction Cards and a Screening Assessment Record Form, four personal ADL tasks and a Neuropsychological Record Form. Studies have been undertaken to evaluate the psychometric properties and clinical utility of the SOTOF Results indicate that the SOTOF is a valid, reliable and clinically useful tool that provides information regarding the relationship between an individual's neurological deficits and ADL performance. The SOTOF identifies information related to four different levels of functioning:

- The person's residual occupational performance in the domain of ADL (by examining ability to perform simple ADL tasks such as feeding and dressing);

- The person's residual and deficit skills and abilities within ADL performance (for example reaching, scanning, grasping and sequencing);
- The performance components (perceptual, cognitive, motor .and sensory) that have been affected;
- The specific neuropsychological deficits (for example apraxia, agnosia, aphasia, spasticity and memory deficits).

Each of these four objectives relate to one of the levels of disability, functional limitation, impairment and pathophysiology (NCMRR, 1992) and address one of four different types of assessment questions (Figure 1.1):

- **How** does the client perform ADL tasks, independently or dependently?
- **What** skills and abilities does the client have intact, and what skills and abilities have been affected by the neurological damage?
- **Which** of the perceptual, cognitive, motor and sensory performance components have been affected by the neurological damage?
- **Why** is function impaired? (Identification of cause through the naming of the specific neurological deficits and underlying pathology.)

The SOTOF is a descriptive test that can also be used to evaluate changes in function over time, as it has been shown to possess good levels of test-retest reliability. Data generated from the administration of SOTOF provides a comprehensive base line from which to set goals, plan treatment and, if appropriate, educate the person and his / her family members / carers.

Dynamic assessment

Dynamic assessment uses an interactive procedure to observe and measure changes in a person's behaviour or function that occur in response to mediation strategies, cues, feedback or task conditions (Haywood & Lidz, 2007; Katz & Bar Haim Erez et al., 2012; Toglia & Cermak, 2009;). The roots of dynamic assessment are based on Vygotsky's (1978) concept of the zone of proximal development (ZPD). ZPD is the gap between what a person has already mastered (the actual level of development/ function) and what he/she can achieve when provided with a supportive learning environment (development/ rehabilitation potential). It has been referred to as the zone of rehabilitation (Cicerone & Tupper, 1986); this 'zone is hypothesized to reflect the clients' region of potential restoration of function or degree of cognitive plasticity" (Toglia & Cermak, 2009, p.570-571). Dynamic assessments focus on process and how performance can be improved owing to some form of guidance (Hadas-Lidor, Weiss & Kozulin, 2011; Sternberg & Grigorenko, 2002). In comparison, static assessments focus primarily on identifying the level of deficit and / or ability (Haywood & Lidz, 2007). Although static assessments can be useful for diagnosis, monitoring progress and discharge planning, their ability to support intervention planning is restricted (Toglia, Golisz & Goverover, 2008).

The SOTOF has a dynamic assessment element which draws on a diagnostic reasoning process (Rogers &Holm, 1991). Although Laver and Powell (1995) never referred to SOTOF as a 'dynamic assessment' in the original test manual, Golisz and Toglia (2003) discussed the SOTOF as an example of a dynamic assessment tool. Therapists administering the SOTOF use prompts and cues, which Toglia (2011) recognises as an important element of dynamic assessment.

Critical appraisal of SOTOF

Literature has suggested the SOTOF is beneficial to use as an assessment of body function and structure for occupational therapists, specifically within neurology and for older people (Clarke et al., 2001; College of Occupational Therapists (COT), 2003, 2004). Letts and Bosch (2001) provided an in-depth critique of the SOTOF, highlighting the strong link to occupational therapy theory and clinical reasoning processes, SOTOF's usefulness for adults with neurological impairments and its ability to evaluate activities of daily living skills. However, they stated that the SOTOF is not particularly useful if the clinician needed to assess all areas of activities of daily living as the SOTOF's four assessment tasks (eating, washing, pouring and drinking, and dressing) are standardised. Letts and Bosch (2001) critiqued the evidence base related to SOTOF's internal consistency, inter-rater reliability, test-retest reliability, content validity, convergent and construct validity, face validity and clinical utility and indicated each of SOTOF's evaluated psychometric properties were good or acceptable. But they noted responsiveness to change had not been examined. Although acceptable levels for test re-test reliability were established (Laver, 1994), the original version of SOTOF is not easy to apply as an outcome measure because it does not provide an overall score. To evaluate clients' performance over time therapists had to examine changes in individual SOTOF items; rather than having a total score that reflects the person's overall test performance. The addition of a six-point ordinal scoring system has been developed for the SOTOF 2nd edition in order to overcome this limitation. Douglas, Letts and Liu (2008) rated the levels of SOTOF's reliability and internal consistency as adequate and noted that advantages of SOTOF were: the low cost; that it can be used as an initial screening assessment because of the short administration time; and that it can be used 'before the client is able to mobilize' (p.24). Law (1997) highlighted the need for further use of SOTOF by clinicians and research to extend the evaluation of SOTOF's reliability and clinical utility.

In a review of SOTOF, McArthur and Spalding (1997, p.501) commented that 'it is evident that SOTOF has been developed after extensive research' and 'the assessor is presented with a high standard of information, record sheets and cue cards to assist in the administration process'. They considered SOTOF a useful 'standardised assessment of neuropsychological deficits of elderly clients' that used 'basic ADL tasks which are familiar to both client and therapist', required 'minimal equipment' and was 'comprehensive in its information' (p.501). McArthur and Spalding (1997) reported that the 'comprehensive glossary, reference list and a list of further assessments' might assist therapists to 'gain a full understanding of the implications of the assessment results' (p.501). However, they noted a number of aspects that could be improved including: an example of a completed record sheet for novice assessors; a summary of the person's occupational performance being placed at the end of each section; a reminder, placed on the SOTOF instruction cards, that tests can be timed for comparison with normative data; and that, although for many test items additional prompts were provided, these were 'not consistent across the assessment format' (p.501). The development of the six-level graduated mediation protocol with specific examples for levels 1 -4 for each SOTOF test item has addressed this identified lack of consistency in the original SOTOF version.

Appraisal of SOTOF materials (Laver & Powell, 1995) identified that it had the potential to be developed to improve the dynamic aspect of the assessment. For example, on the SOTOF instruction cards, the 'additional prompts required' section was combined with 'further assessments required' but these sections were mainly focused on suggested assessments and tests. When there was guidance for prompting provided it lacked detail and relied on the clinicians' clinical reasoning, a potential challenge for novice therapists, and there was no grading element to the prompts suggested. Appraisal of SOTOF indicated there was potential for further development of SOTOF's dynamic element. This led to a further content validity study to inform the enhancement of the dynamic assessment phase and the development of this SOTOF 2nd edition.

Figure 1.1: The SOTOF assessment model – the relationship between occupational performance and neurological deficits

DISABILITY	FUNCTIONAL LIMITATION	IMPAIRMENT	PATHOPHYSIOLOGY
Definition:	Definition:	Definition:	Definition:
Inability or limitation in performing socially-defined activities and roles within a social and physical environment as a result of internal or external	Restriction or lack of ability to perform an action or activity in the manner or range considered normal that results from impairment.	Loss and/or abnormality of mental, emotional, physiological or anatomical structure or function; including secondary losses and pain.	Interruption or interference of normal physiological and developmental processes or structures.
OCCUPATIONAL PERFORMANCE	SPECIFIC SKILL AND ABILITY	PERFOMANCE COMPONENTS	NEUROLOGICAL DEFICIT
Personal ADL	**Skills**	**Components**	**Diagnosis**
• Eating • Washing • Pouring and Drinking • Dressing	• Reaching • Scanning • Naming • Sequencing	• Perceptual • Cognitive • Motor • Sensory	• Apraxia • Dysphasia • Agnosia • Spasticity
HOW?	**WHAT?**	**WHICH?**	**WHY?**

Laver and Baum (1992). Note: Definitions of disability, functional limitation, impairment and pathology are modified from work by the NCMRR (1992)

2 Test Administration and Scoring

General introduction to the content and administration of the SOTOF Test environment, materials and equipment

Each of the four **ADL tasks** is then administered using the standardised instructions **Please read this section before attempting to administer the SOTOF**
The SOTOF involves the administration of five sub-tests: a Screening Assessment; and four tests based on the observation of simple personal ADL tasks (eating, washing, pouring and drinking, and dressing). An Instruction Card is provided for each of the five sub-tests. Instruction Cards are used to guide test administration, score observed behaviour and consider the underlying neuropsychological function. The person's performance is recorded directly on to Record Forms during testing, which provide a profile of the person's skills and deficits. Record Forms are reviewed and results from all five sub-tests are drawn together and summarised on a Neuropsychological Checklist at the end of testing. This provides a profile of neuropsychological function and occupational performance in the four ADL areas.

The Screening Assessment is administered to identify the person's basic level of functioning and to determine whether the client meets the criteria needed to attempt the ADL tasks. If you already know your client well enough to establish his / her suitability to be assessed on SOTOF then you do not need to administer the screening assessment.

Instructions for administering the SOTOF are provided in the 1st (left-hand column) of the **SOTOF Instruction cards** (see Figure 2.1). The person's performance is recorded on the Record Form initially using a simple categorical two-point scoring system (able or unable). The Instruction Cards list the main deficits associated with an inability to perform each test item in the 2nd column. If the person has been unable to do a test item then the assessor applies the graduated prompt protocol, examples for level s 1 to 4 for each SOTOF test item are provided in the Instruction cards in the 3rd column. The 4th column (on the right-hand side) of the Instruction cards briefly lists other assessments which may be used if further clarification of a deficit is required. Detailed scoring guidelines and examples are given for each potential deficit listed on the ADL and Neuropsychological Record Form later in this manual. A Glossary is provided to explain the terminology used to label deficits. In addition, sources of further information about deficit areas and references for other neuropsychological assessments are provided in the Further Reading section at the end of this manual.

Test environment
The SOTOF may be administered in any suitable clinical or home setting. To maximise the person's performance try keep noise and visual distractions to a minimum, provide good lighting, maintain a comfortable temperature and restrict draughts.

Equipment and materials: The SOTOF requires no special test materials and equipment. Those required are everyday items, easily available in many clinical settings or in the person's own home. Items may, however, be purchased and kept together as a kit for the assessment. A summary of the equipment and materials required to administer the SOTOF can be found in Table 2.1. Ecological validity is improved if the person's own belongings are used and if he / she is given a choice in the selection of the food and drink to be consumed and the garment to be worn. Furniture may be the client's own, or from the hospital ward or department. It should be an appropriate height and type for the person's height and build, for example people are likely to gain most support from a sturdy chair with back support and arms. If the items are being purchased to make up a kit, factors such as weight, fragility and colour should be considered. People with motor deficits may have difficulty manipulating and lifting heavy items. Colour vision alters with ageing owing to yellowing of the lens. Dark colours, particularly at the blue-green end of the spectrum, are the most difficult for older adults to perceive. Equipment should ideally be lightweight, made of unbreakable substance (for example plastic or melamine) and of different plain, bright primary colours. Adaptive equipment (such as lightweight, large-handled spoon) may be used, but its use should be noted on the Record Form.

Table 2.1: Equipment and materials required for SOTOF

Sub-test	Furniture	Equipment	Materials (consumables)
Screening Assessment	Table and 2 chairs	• Cup • Pen	• None required
Eating task	Table and 2 chairs	• Bowl • Non-slip mat • Spoon	• Food
Washing task	Table and 2 chairs	• Washing bowl • Hand towel • Non-slip mat	• Warm water to ¾ fill washing bowl • Soap
Pouring and Drinking task	Table and 2 chairs	• Jug • Cup • Non-slip mat	• Cold drink to ½ fill jug
Dressing task	Table, chair for tester, bed, plinth or chair for client.	• Front fastening (buttons or zip) long-sleeved garment such as a shirt, blouse, cardigan or jacket of suitable size and type for client. • Large bright coloured button.	• None required

Food for '**Task 1: Eating**' should be soft and ideally in small pieces which do not require cutting. Please ask the person about any history of diabetes or food allergies before selecting a suitable food choice. Suitable food includes: tinned or fresh fruit–salad; small pieces of fruit in natural juice; cereal with milk; soup with small pieces of vegetable or meat; or mince in gravy. The food selected ideally should be relevant to the time of day when the assessment is being undertaken.

Water should be warm for '**Task 2: Washing and Drying hands'**. Please test the temperature of the water yourself before the assessment of Task 2, as the person may have a sensory deficit and be unable to discriminate hot, burning water. Some people may be allergic to scented soap, so ask the person about any skin allergies before the assessment is undertaken. Ideally, use a non-scented allergy tested soap or the person's own soap. This can either be a bar of soap, or if the person prefers liquid soap in a dispenser.

For '**Task 3: Pouring and Drinking'** a cold drink is required; suitable drinks include: water; fruit juice; squash; and milk.

For '**Task 4: Dressing**' it is preferable to use the person's own garment. This needs to be a front fastening garment with buttons, for example: cardigan; jacket; shirt; blouse; or dressing-gown / bed jacket.

Important note: Take into account diabetes or food allergies when selecting food and drink, and skin allergies when selecting the soap.

Using the SOTOF Instruction Cards

An example of a Screening Assessment Instruction Card is given in Figure 2.1. Instruction Cards are provided for the Screening Assessment and the four ADL tasks. The Instruction Cards have been developed to structure and prompt your observations and interpretation of the person's observed behavioural responses. Four types of information are provided in the Instruction Cards:

1. instructions for administering the test (standardised assessment phase);
2. possible diagnoses provided to facilitate hypothesis generation (aids diagnostic reasoning);
3. examples of mediation for levels 1-4 of the graduated mediation protocol (dynamic assessment phase);
4. instructions to guide the acquisition of additional information about the client's function (e.g. examples of further assessments which might be useful).

These four types of information are provided for each SOTOF test item and the Instruction Cards are, therefore, divided into four columns headed: Task and instruction; Possible area of deficit; Graduated mediation protocol examples; and Further suggested assessment.

Task and instruction

Instructions for administration include: verbal instructions given to the client to elicit the performance to be observed (such as: *'Put the spoon on the table on the right of the bowl'*); instructions for the tester, (such as: Put spoon in hand on the opposite side to the cerebral lesion - if fails to identify, try with the other hand); and prompts regarding aspects of performance which the tester should note (such as: note if client scans table for objects).

Figure 2.1: Start of the Screening Assessment Instruction Card (showing the four types of information columns).

	COLUMN A	COLUMN B	COLUMN C	COLUMN D
	Screening Assessment **Task and instruction** **(Standardised assessment phase)**	**Possible area of deficit** **(Aids hypothesis generation)**	**Graduated mediation protocol examples** **(Dynamic assessment phase)**	**Further suggested assessment** **(To aid diagnostic reasoning and test hypotheses further)**
1.	(EL) **Ask:** *'What is your name?'*	• Hearing • Language comprehension and/or expression • Orientation	1. General verbal cue: '*Can you tell me your name out loud?' 'What are you called?'* 2. Gestural Cue: N/A 3. Specific feedback: *'My name is … What is your name?'* 4. Physical Assistance/modifications: N/A	Check if client uses hearing aid. Test hearing, e.g. Free field hearing test. Test comprehension / expression. Test orientation, e.g. CAPE
	Start by following the instructions in column A to administer the standardised element of the SOTOF. *Instructions in italics tell you what to say.* **Note whether the person is able or unable to perform this test item following this instruction.** **If Able – move to the next item below in column A** **If Unable - move to column B**	**This column advises you of the more common reasons a person may have been unable to successfully complete the item.** **It is helpful to have these possible areas of deficit in mind for the dynamic assessment element.** **Now move to column C**	**This column gives you example instructions for administering the dynamic assessment for this SOTOF test item. Suggestions for mediation for levels 1 - 4 are provided. Start at level 1. Give 2 instructions at each level before moving to the next level. If the person responds note the level of mediation that was successful. Then go to the next item below in column A** **If they do not respond to any mediation, do the item for the person. Now move to column D for further ideas**	**This column provides suggestions for further assessments and / or other things to observe and note down.** **This information can be useful when completing the SOTOF Scoring form and deciding if any further assessment is required.** **Then go to the next item below in column A**

Possible area of deficit

Potential hypotheses (neuropsychological deficits) are given to explain dysfunctional cues observed in each of the test items. The diagnoses are listed in the 2nd column of the Instruction Card alongside the associated instructions for that test item. The tester should review each item in the light of observed behaviour. Possible deficits are recorded on the record forms in the column headed 'Hypotheses, cues and comments'. Further information about these deficits can be obtained in the Glossary (pages129-133), Guidelines for observing and scoring the Neuropsychological items (pages 34-39), and Further Reading and Other Assessments at the end of this manual (pages134-135). These resources are designed to facilitate decision-making during the test and serve as a teaching tool enabling the therapist to build on his or her knowledge.

Applying the SOTOF Graduated mediation protocol: Dynamic assessment

The SOTOF six-level graduated mediation protocol is outlined below in Table 2.2. If the person is able to perform the SOTOF test item independently then the therapist places a tick in the 'able' column on the Record form (Figure 2.2) and then ticks the box to indicate a score of 0 = independent. If the person is unable to complete the test item independently then the dynamic assessment phase is initiated. The assessor (clinician / therapist) must provide the mediation (e.g. prompts, cues, assistance, adaptation) in the order provided in graduated mediation protocol, starting with the least intrusive (level 1 mediation) first. The clinician should allow the person adequate time to respond to mediation and must give two forms of mediation (e.g. cues, prompts) at each level of the graduated mediation protocol, before moving to the next, higher, level (Baum and Wolf, 2013). The clinician must ensure the overall ADL task is completed successfully, even if this requires the highest level of the graduated mediation protocol, 'do for the person' (Baum and Wolf, 2013) This is because it is an interactive procedure and will contribute to maintaining the motivation for both yourself and the client.

When using the record form tick the highest level of the graduated mediation protocol carried out in each subtest to complete the task. In the summary section of each task the clinician should comment on the learning potential of the person and how effective the mediation methods (prompts / cues / modifications / assistance) were. The clinician should also comment on which graduated mediation methods were the most effective for that individual, as this could inform future assessments and/or interventions. The higher the score the more assistance is required by the person. To complete the final scoring in the neuropsychological checklist the clinician should look down all the scores within each task and whichever sub-test item scores the highest on the graduated prompt protocol is the one recorded for that task. This is because somewhere within the task the person needed that level of assistance to be successful. Examples of mediation methods (prompts / cues /modifications / assistance) for levels 1 to 4 for each sub-test item can be found in the third column of the Instruction Cards (see Figure 2.1). Unless they are not applicable for that type of sub-test item, for example, if the person has their eyes closed to offer a gestural cue is not appropriate. As level four has a variety of different mediation options for the clinician to use, when completing the record form, the specific type of prompt / cue / assistance / modification provided at this level should be noted on the form.

Table 2.2: SOTOF Graduated mediation protocol

0	**Independent**	The person is independent completing the task. No prompting or assistance is required from the clinician.
1	**General verbal cue**	This could be a statement (Katz et al., 2011) e.g. 'take your time' or could be a general question e.g. 'what do you think is the next step?' or 'what else might you need to complete this task?' (Baum and Wolf, 2013 p.3). This is not an action or telling the person what to do.
2	**Gestural Cue**	This could be miming the action that is required to complete the particular task or a movement that may guide the participant. This may include pointing to where they might find an item or pointing to equipment they may need to complete the task (Baum and Wolf, 2013).
3	**Specific feedback (verbal cue or prompt)**	This can be a verbal cue, in the form of feedback (Katz et al., 2011) such as 'there is a mistake, can you try and correct it' or a specific verbal prompt command such as 'pick up the cup' (Baum and Wolf, 2013 p.3).
4	**Physical assistance** and / or **Co-active assistance** and / or **Modifications** and / or **Demonstration**	**Physical assistance:** This clinician physically supports the person to complete an action, e.g. hold the shirt whilst the person puts his / her first arm in the sleeve (Baum and Wolf, 2013). **Co-active assistance:** The clinician physically guides the movement but allowing the person to lead and withdraws the physical assistance if the person takes over the movement (Sanderson and Gitsham, 1991). **Modifications:** The clinician reduces the amount of stimuli or modifies the environment to reduce the task demand (e.g. changing the physical environment; Katz et al., 2011). **Demonstration:** The clinician may also do the action using task items in order for the person to copy (Katz et al., 2011). The person should still be attending to the task (Baum and Wolf, 2013).
5	**Do for the person**	The person is unable to complete the task so the clinician completes the task, or the part of the task, for the person.

Further prompts/assessment

Instructions are provided to guide further data collection for occasions where there is a differential diagnosis. For example, failure to identify an object could be the result of several deficits: failure to comprehend the instructions; failure to see the object; inability to express the name of the object or indicate the) object; or inability to recognize the object seen. To clarify the area of deficit several further assessments or prompts may be required to assess function in the areas of comprehension, vision and expressive language. In this example, if the client understands the instructions, can see the object, can either name or indicate objects and is willing to cooperate, then failure to identify the object is most likely to be the result of a visual agnosia. Information outlining prompts, areas for further assessment and specific instructions to elicit further performance are provided in the right-hand column of the Instruction Card alongside the associated diagnoses. Examples of information provided in this column include: 'Assess for visual field loss . . . Point to the right of the bowl and ask the client to place the spoon there". Ask client to describe the task. Retest (R) and (L) with other objects and body parts.'

References for further assessments

When further assessment is required to clarify function some assessments are suggested in the 4th column of the instruction cards and at the end of this manual. These assessments were selected following consultation with geriatricians, speech therapists, ophthalmologists, audiometricians and psychologists. Most of these assessments are standardised, widely known, quick and simple to administer and score, and are relevant for use with an older client group.

The Screening Assessment

The Screening Assessment identifies the client's basic level of function. The essential behaviours required to complete the four ADL tasks were identified through activity analysis (See Table 2.3). These behaviours serve as four criteria which clients are required to meet if SOTOF is to be administered in its standardised form:

1. The client should be able to comprehend verbal, written or demonstrated instructions (if she does not speak English, or if her speech is difficult for the therapist to comprehend but is understood by the client's carer or speech therapist, an interpreter may be used).
2. The client should be able to see objects placed on a table up to 45" from the client in the mid-line of her visual field (glasses should be worn if the client usually wears glasses during ADL tasks).
3. The client should have gross functional use of one upper limb, sufficient to lift and manipulate test materials.
4. The client should be able to sit upright in a chair for the anticipated duration of the test; support cushioning may be used to assist sitting balance (see Appendix for normative standards for administration time).

Administering the Screening Assessment

The Screening Assessment is administered to the client prior to undertaking the four ADL tasks. If the client has been assessed previously and is known to meet the four testing criteria then testing may begin with the four ADL tasks. The Screening Assessment Record Form is used to record the client's hand dominance

and information about equipment required by the client during performance of the ADL tasks (for example glasses, hearing aid, wheelchair and support cushioning). The Screening Assessment items on the Instruction Card are based on occupational therapy practice and involve the evaluation of gross vision, orientation, language, and trunk and upper limb motor function. If the client is known to fulfil the basic criteria from information previously gathered, the Screening Assessment may be omitted. If the client fails, or is known not to be able to perform any of the Screening Assessment items (numbers 1 to 5), further assessment of these areas should be conducted prior to administration of the ADL tasks. If the deficit can be improved with the use of equipment (for example glasses, hearing aid, cushioning), alteration of instruction format (verbal, written, demonstrated) and/or the use of one-handed techniques, these should be implemented. Any equipment used and/or changes made should be noted in the summary- section of the Screening Assessment Record Form. If the client is unable to fulfil the basic criteria (is unable to comprehend instructions) some items on the ADL tasks may provide useful information; for example, in the Eating task hand the client a spoon and bowl of food and observe the response.

Scoring the SOTOF

An example of a SOTOF Record Form is given in Figure 2.2. The Instruction Cards and Record Forms follow the same order of items. Items are numbered to help you to match items on the Record Forms and Instruction Cards. Observations are recorded on the Record Form using three different methods: a nominal, two-category scoring system (able / unable); an ordinal six-point scoring system which aligns to the six-level graduated mediation protocol; and a descriptive, qualitative, format.

Category scoring system (standardised assessment): The client's performance for each SOTOF item is first recorded using a two-category nominal scoring system by placing a tick in the relevant column headed either ‘Able (yes)’ and 'Unable (no)' on the Record Form.

Six-level ordinal score (dynamic assessment): If the person has been able to do the test item then score as 0 = independent. If they have been unable to do the test item independently, the therapist follows the graduated mediation protocol (see Table 2.2) and scores accordingly: 1 = general verbal cue; 2 = gestural cue; 3 = specific feedback or prompt; 4 = Physical assistance, and / or Co-active assistance and / or Modifications; or 5 = do for the person.

Descriptive recording (diagnostic reasoning): Space is provided against each SOTOF test item for notes regarding the level(s) of mediation provided, related hypotheses regarding deficits (cognitive, perceptual, motor and / or sensory) and suggestions for any additional testing required. Use the summary section on the Record Forms to make descriptive statements about the client's ability to perform each ADL task and useful qualitative information, for example: the client's manner during testing; the test environment; any interruptions or distractions occurring during the assessment; client’s report of fatigue and / or pain which might have affected his / her performance on the assessment. If the client is able to communicate, further informal questioning may be used after the assessment to clarify the client's experience of the problem and is recorded in the summary section.

Table 2.3: Summary of research studies to evaluate the psychometric properties and clinical utility of SOTOF

Study	Purpose	Research questions	Design	Subjects	Results	Contribution to instrument development
Content validation Activity analysis (1990). Funded by South West Thames LORS Grant: April 1990-91.	Provide content validation for SOTOF based on identification of behaviours, elicited by the four ADL tasks that represent the constructs to be evaluated by SOTOF.	What behaviours are elicited by the four ADL tasks? Which behaviours relate to which of the constructs to be evaluated by SOTOF?	Detailed analysis of the four ADL tasks defined as the test domain. Descriptions of four ADL tasks and a structured activity analysis format given to experienced OTs who performed and analysed the tasks.	Six OTs with experience in gerontology and neurology who were working with stroke patients in two health authorities in south-west London.	A qualitative analysis was undertaken to compare the six activity analyses and showed considerable consensus. Data identified the component steps of each task, input received from the materials and performance of each task, the expected behavioural responses to each task step and the nature of feedback.	Third stage of establishment of content validity: information obtained from analyses formed the basis of the content of the SOTOF protocols and forms. (Studies that provided the first two stages of content validity are summarized in Table 3.1 on p. 22.)
Content validation Expert and peer opinion - local survey through interview (1991).	Provide content validation for SOTOF based on the opinion of peers and experts in the field of occupational therapy.	Does the SOTOF possess content validity according to the clinicians who will administer the SOTOF and experts?	SOTOF manual provided to peer OTs who administered the test and to experts who reviewed test materials. Feedback collected through interviews	Five OTs working in gerontology in south-west London. Two OT experts based at two National Rehabilitation Centres.	A qualitative analysis of feedback obtained through unstructured face to face and telephone interviews indicated that the SOTOF was perceived by peers and experts to possess content validity	Fourth stage of establishment of content validity.
Content validation Peer opinion national survey (1991). Funded by South West Thames LORS Grant: April 1990-91.	Provide content validation for SOTOF based on the opinion of peers in the field of occupational therapy.	Does the SOTOF possess content validity according to the clinicians who will administer the test?	SOTOF manual and a structured feedback questionnaire provided in a postal survey to peer OTs who administered and reviewed the test.	Forty-four qualified OTs working across the United Kingdom (UK), Ireland and Belgium.	SOTOF was perceived by peers to possess content validity. Four constructs, perceptual, sensory (touch, vision and hearing), motor and cognitive (memory and language), emerged as the content base of SOTOF along with ADL performance.	Final stage of establishment of content validity.

Figure 2.2: The first part of the Screening Assessment record form

Key: **(EL)** items can be administered to clients with expressive language

(ED) items provide alternative assessment methods for clients with expressive dysphasia

Client's name:	**Date:**
Tester's name:	

	Item	Able	Unable	Level of mediation required	Hypotheses, further assessments required, comments
1	Name			0. Independent ☐ 1. General prompt ☐ 2. Gestural cue ☐ 3. Specific feedback/cue ☐ 4. Physical assistance ☐ 5. Do for client ☐	
2	Vision			0. Independent ☐ 1. General prompt ☐ 2. Gestural cue ☐ 3. Specific feedback/cue ☐ 4. Physical assistance ☐ 5. Do for client ☐	
3	Sitting Balance			0. Independent ☐ 1. General prompt ☐ 2. Gestural cue ☐ 3. Specific feedback/cue ☐ 4. Physical assistance ☐ 5. Do for client ☐	
4	Upper limb	☐ Right ☐ Left	☐ Right ☐ Left	0. Independent ☐ 1. General prompt ☐ 2. Gestural cue ☐ 3. Specific feedback/cue ☐ 4. Physical assistance ☐ 5. Do for client ☐	
5	Hand grip	☐ Right ☐ Left	☐ Right ☐ Left	0. Independent ☐ 1. General prompt ☐ 2. Gestural cue ☐ 3. Specific feedback/cue ☐ 4. Physical assistance ☐ 5. Do for client ☐	

The Record Forms are headed with spaces to record details about the client, any equipment used and tester. Observation of the client's responses should be recorded on the ADL and Neuropsychological Record Form either during or immediately after testing. Performance of each of the behavioural task components is identified through a numbered test item on the Instruction Card and Record Form and is recorded separately. The components are summarised descriptively on the Record Form, for example 'Identifies soap through touch'.

Scoring example related to the Screening Assessment
The first component on the Screening Assessment requires the tester to ask the client's name. If the client has failed to respond, the 'Unable (no)' box will be ticked. The tester then refers to the Instruction Card to identify possible deficits and any further assessment. For example: the tester will need to consider whether the client can hear; whether the client can comprehend; whether the client is orientated for person (knows personal details, including her own name); and whether the client's expressive language is intact. If, for example, on further testing, the client is able to read and respond correctly to the written instruction, hearing will have been identified as the deficit. Therefore, the tester will tick the 'Hearing acuity' deficit box in the 'Screening Assessment' column on the Neuropsychological Checklist. In which case, the large written instructions would then be used for the administration of the Four ADL Tasks.

The four SOTOF ADL tasks

Description of the four ADL tasks

The SOTOF personal ADL tasks are suitable for people of any gender and from various cultural backgrounds. They can be performed one-handed, using a non-dominant hand (hemiplegic techniques may be used, for example for dressing). The client can be seated. Each ADL task is administered following a standardised Instruction Card; the client's response is observed, scored, and recorded on a Record Form for that task.

Task 1 - Eating task: Eating from a bowl using a spoon
The person is seated in a suitable chair at a table. A bowl of food is eaten using a spoon. The type of food selected for testing can be based on the client's preference, culture and religious requirements. The client should be offered a choice of food; it may be helpful to negotiate food prior to testing to ensure adequate preparation. Swallowing may sometimes be affected following a stroke. Liquid, dry foods and large pieces of solid food are more difficult to swallow than semi-solids and small pieces of food in liquid. Suggestions for suitable foods for this task include small pieces of soft fruit in juice, cereal with milk, mince in gravy and rice dishes. The therapist should be aware of any diabetes or food allergies. A non-slip mat should be placed under the bowl to prevent slipping and facilitate one-handed eating, if necessary owing to hemiplegia. A large-handled spoon may be used to facilitate grip (for example for people with arthritis). It is preferable to use a plastic or melamine bowl as these are lighter to lift and less likely to break if dropped.

Task 2 - Washing task: Washing hands in a bowl of water

The person is seated in a suitable chair at a table. Hands are washed in a bowl of warm water using soap and are dried on a hand towel. The bowl is placed on a non-slip mat to facilitate one-handed washing, if necessary, owing to hemiplegia. The size of bar of soap should not be too small to grasp or liquid soap in a dispenser may be used if the person prefers / usually uses liquid soap. Some clients may be allergic to scented soap, so ask about any skin allergies prior to testing. The person's own soap, towel and bowl can be used if possible.

Task 3 - Pouring and Drinking task: Pouring liquid from a jug into a cup and drinking from the cup

The person is to be seated in a suitable chair at a table. The type of drink selected can be based on the client's preference, culture, and religious requirements. A choice of drink should be offered; it may be helpful to negotiate the drink to be used prior to testing to ensure adequate preparation. The therapist should be aware that swallowing may sometimes be affected following a stroke. Liquids of extreme temperatures (boiling hot and freezing cold) can be more difficult to swallow than cool or warm liquids. Suggestions for suitable drinks include water, fruit juice, squash, and milk. The therapist should ask the person about diabetes or food allergies prior to testing. The person is required to pour a drink from the jug into a cup, and then drink from the cup. The type of cup can be varied depending on the needs of the client, for example a large-handled, lightweight cup for clients with arthritis. Positions of objects, as identified in the relevant Instruction Card, may be reversed if the person is left-handed. It is preferable to use a plastic or melamine jug and cup as these are lighter for the person to lift and less likely to break if dropped. The person's own utensils can be used if possible. In in-patient ward settings, many patients are provided with a plastic jug and beaker by their bedside.

Task 4 - Dressing task: Putting on a long-sleeved, front-fastening upper garment

The person's own clothing should be used where possible, and he / she should be encouraged to choose a garment to wear. The person puts on a long-sleeved, front-fastening garment. The garment should be the appropriate size, style, and type for the person. The person may choose to use dressing aids, such as dressing sticks and button hooks, and the hemiplegic dressing method, if necessary, in which case the therapist should note the aids and / or method used in the summary section of the dressing task record form.

Description of ADL task test items

The SOTOF ADL assessment tasks include three types of test items:

1. Items which involve the observation of discrete behavioural components drawn from the normal performance of each ADL task.
2. Items which involve a question-response format in which the client is asked to name or indicate objects and colours, and to describe the use of objects, the temperature of water and the taste of food and drink.
3. Items using materials / objects used to perform the ADL tasks, for example the identification of objects through touch, and miming the use of objects on command.

Most constructs addressed by SOTOF are evaluated through observation of behaviour elicited by the performance of the ADL tasks. This behaviour is prompted by the simple commands outlined in the Instruction Cards, such as *'Eat the food in the bowl using the spoon'*. Each discrete behavioural component has been identified through activity analysis and forms a separate test item. For example, in the Eating task, items include 'Reaches for spoon', 'Takes food into mouth', 'Replaces spoon in bowl' and 'Repeats sequence'.

Some constructs, although components of the tasks, can only be elicited specifically using additional commands. For example, discrimination of temperature is a component of the washing task but cannot be tested through observation alone; the person is, therefore, questioned about his / her perception of the temperature of the water. Taste of both the food and drink, and recognition of the name, use and colour of items, are also assessed through questioning. For people with expressive dysphasia some of these items can be assessed using alternative instructions that require the person to point at the correct object or colour. These alternative instructions are marked with ED (for expressive dysphasia) on the SOTOF instruction cards.

A few deficits cannot be addressed as part of the usual performance of these ADL tasks but are important areas to evaluate during comprehensive neuropsychological assessment. These constructs include tactile discrimination, right/left discrimination and aspects of apraxia. Additional test items have been constructed using the equipment and materials from each ADL task. These test items have been based on accepted testing methods as identified through a literature review and critique of other perceptual and cognitive assessments.
In everyday life tactile discrimination is required when a client performs an action when the hands are not within his / her visual field. For example, fastening buttons on the back of clothes, or searching for items in the dark. In these instances the client has a concept of what she should be feeling. In daily life the client usually has cues from the environmental context which assist the identification of objects through touch.

Tactile discrimination is usually assessed by presenting clients with a series of objects whilst their vision is occluded. Several variations of this form of assessment have been described by Zoltan *et al* (1986). An example of this form of testing can be found in the Chessington OT Neurological Assessment Battery (COTNAB) (Tyerman et al*.,* 1986). The SOTOF contains a tactile discrimination item for each of the four ADL tasks. These items involve the naming of common objects drawn from the task following identification through touch. In order that the object is not handled by the client prior to the test item, the tactile discrimination items are placed at the beginning of each of the ADL Instruction Cards. Some of the items from the ADL task are laid out on the table in front of the client to set the context for the test, one item is kept aside and then handed to the client for identification once he / she has closed his / her eyes. The objects used (spoon, soap, cup and button) provide a range of sizes and textures.

Right/left discrimination and **spatial relations** are usually evaluated by asking the client to indicate parts of objects or environments, or to place objects in relation to each other on command. Some tests involve the self-identification of

body parts on command. During the performance of the ADL tasks there are several occasions when the client replaces objects on the table. These occasions are used to give the client specific commands regarding the positioning of the object. After the tactile discrimination item clients are left holding an object, and are then instructed to place the object on the right (or left) of one of the objects already on the table. For example, in the Eating task clients are instructed to 'put the spoon on the table on the left of the bowl'.

Apraxia is a complex construct, defined as the 'inability to perform motor activities although sensory motor function is intact and the individual understands the requirements of the task' (Crepeau, Cohn and Boyt Schell, 2009, p. 1154). Apraxia has been sub-divided into several specific types which are usually referred to with a clarifying descriptor. Types of apraxia identified through occupational therapy and neuropsychological assessments include: constructional apraxia; dressing apraxia; motor apraxia; ideomotor apraxia; and ideational apraxia. These last three types of apraxia are very similar and difficult to differentiate, so it is not always possible to discriminate between these types simply by observing the performance of ADL tasks. Therefore, in addition to the observation of real object use, testing usually involves the performance of gestures, miming the use of objects and imitation of non-representational (or nonsense) movements (Edwards et al*.,* 1991). Several SOTOF items are used to differentiate between these different types of apraxia. These items involve the client describing, miming and then demonstrating the use of the objects, prior to actual object use.

Formats of instructions that can be selected for specific client groups

Hearing acuity may be reduced as a consequence of the normal ageing process. If the client has a hearing deficit, written rather than verbal instructions might be preferred. Visual acuity may also be affected by primary ageing. A set of written instructions are provided in clear, enlarged text (see pages 111-128). Some people, for example with dementia, benefit from additional written instructions. These instructions may be photocopied / printed as required. If you work with people with dementia or with hearing deficit frequently, then we advise that you print and laminate a set of the enlarged instruction cards.

Some clients experience **expressive aphasia** or **dysphasia** as a result of conditions such as stroke. Alternative forms of testing, such as pointing to named items and colours, have been provided where possible. These alternative forms of testing are given in the Instruction Cards and are coded with the letters (ED) for easy identification. The need for these alternative forms of testing should be identified during the administration of the Screening Assessment. Receptive aphasia may result from a stroke. If it is identified in advance, omit the instructional test items and present the client with the visual and tactile cues provided by the test materials. For example, hand a bowl of food and a spoon to the client and observe his / her response.

Administration of the ADL tasks

The ADL tasks can be given in any order but have been ordered in general order of difficulty and common recovery of function (i.e. feeding, washing hands, pouring drink, dressing). Ideally the tasks should be given at an appropriate time of day. The tasks can be administered together or individually. If the person tires quickly then a break should be given between tasks. The whole assessment should be administered within three days to prevent a possible difference in performance due to recovery. Ideally, the tasks should be given on the same day. A separate Instruction Card is provided for the Screening Assessment and each of the four ADL tasks. Each of these five subtests is administered by following the instructions in the first column on the relevant Instruction Card.

Scoring the ADL tasks

The ADL tasks are scored on the Record Forms in the same manner as the Screening Assessment (see page 24 for instructions regarding the SOTOF scoring method and recording on the Record Forms). At the end of each of the four ADL tasks add up the scores (6 level ordinal scoring system related to the graduated mediation protocol) given in the 3rd column. The total score is then placed into the 'Total score for graduated mediation' box at the end of each Record form.

Use the summary section on the Record Form to make descriptive statements about the client's ability to perform each self-care task, in addition to qualitative information, for example, about the client's manner during testing, test environment and interruptions to the testing procedure.

The Neuropsychological Checklist and Summary Form

An example of a page from the neuropsychological section of the ADL and Neuropsychological Record Form is given in Figure 2.3. This part of the Record Form identifies the abilities and deficits which are addressed by the SOTOF. Neuropsychological checklist items are grouped under broad headings: language; hearing; cognition; motor; sensation; vision; agnosia; apraxia; body scheme; spatial relations; and perseveration. All terminology used in the assessment is defined in the Glossary (see pages 129-133). The behaviours associated with these items are described in the Guidelines for observing and scoring the Neuropsychological items (See pages 34-39). At the end of each ADL task deficits are recorded on to the Record Form by ticking the appropriate box. When all the tasks are completed the Record Form provides a summary profile of function, so that patterns of deficits can be identified.

Completing the Neuropsychological Checklist

The final stage of the SOTOF involves the evaluation of diagnostic hypotheses about the underlying pathology. A checklist of possible neuropsychological diagnoses is provided to prompt thorough hypothesis generation and review. This third scoring method involves checking the identified neuropsychological diagnoses

against the task in which cues, associated with the related diagnostic hypothesis, were observed. The resulting Neuropsychological Checklist provides a visual summary of function and dysfunction indicating existing function (indicated by a and where a deficit was impacting across tasks or related only to a specific activity. Detailed scoring guidelines for each of the items on the Neuropsychological Checklist are provided on pages. 34-39. Place ticks in the boxes that correspond to the deficits you feel are indicated by the client's performance and the tasks in which the indicative performance was observed. Look down the left-hand column for deficit(s) and across the columns at the top for tasks.

At the end of each of the four ADL tasks add up the scores (6 level ordinal scoring system related to the graduated mediation protocol) given in the 3rd column. The total score is then placed into the '**Summary scores for the four ADL Tasks'** table (see Table 2.4) which is located on the front page of the SOTOF record and scoring forms Then divide each score by the maximum possible score and multiple by 100 to calculate the percentage score. For example, if a person scored 80 on Task 1 – Eating then you would divide 80 by 145 = 0.55 and then multiply this by 100 to get a percentage of 55%.

Then in the **'Overall Score for each of the four ADL tasks'** table which is located on the 2nd page of the SOTOF record and scoring forms (see table 2.5), in the place ticks in the boxes that correspond to the highest level of mediation required for any of the test items in each of the four ADL task.

Table 2.4 Summary scores for the four ADL Tasks

Tasks	Total Score	Percentage score
Screening Assessment	/25	
Task 1: Eating	/ 140	
Task 2: Washing	/ 135	
Task 3: Pouring and Drinking	/ 135	
Task 4: Dressing	/ 100	

Figure 2.3: First part of the Neuropsychological Checklist

To score: Place ticks in the boxes that correspond to the deficits you feel are indicated by the client's performance and the tasks in which the indicative performance was observed. Look down the left-hand column for deficit(s) and across the columns at the top for tasks.

Client's name:	
Tester's name:	**Date of testing:**
Diagnosis:	

DEFICIT	SCREENING ASSESSMENT	EATING TASK 1	WASHING TASK 2	POURING AND DRINKING TASK 3	DRESSING TASK 4
LANGUAGE					
Comprehension					
Expression					
HEARING					
Hearing acuity					
Auditory agnosia					
COGNITION					
Orientation					
Attention					
Short-term memory					
Long-term memory					
Initiation					
MOTOR					
Abnormal tone (spasticity or flaccidity)					

Table 2.5: Overall Score for each of the four ADL tasks

OCCUPATIONAL PERFORMANCE	0 INDEPENDENT	1 NEEDED GENERAL PROMPT	2 NEEDED GESTURAL CUE	3 NEEDED SPECIFIC FEEDBACK / CUE	4 NEEDED PHYSICAL ASSISTANCE	5 DO FOR CLIENT
Eating: Client's ability to eat independently from a bowl.						
Washing: Client's ability to wash and dry hands.						
Pouring and Drinking: Client's ability to pour from a jug and to drink from a cup.						
Dressing: Client's ability to put on a front-fastening, long-sleeved garment.						

Normative Standards

Time

Normative standards were developed on the time taken to perform the Screening Assessment and four ADL tasks. To facilitate comparison with other tests, SOTOF derived time classifications have been constructed through an analysis similar to that undertaken by the authors of COTNAB (Tyerman *et al,* 1986), and by Laver and Huchison (1993) in a study that obtained normative data on the COTNAB for an elderly population. The time taken (in minutes and seconds) on each sub-test is used to produce a derived time grade; this grading system is based upon the standard deviations (sd) and percentile equivalents of the normal curve. On this basis it is possible to measure, in a standard way, the extent to which an individual score deviates from the scores of the normal population.

The SOTOF derived time classifications were calculated using the following percentile equivalents: 0: -3 sd; 2nd: -2 sd; 16th: -1 sd; 50th: mean (0); 84th: +1 sd; 98th: +2 sd; and 100th: +3 sd. The derived score for time used for the SOTOF is classified in terms of five categories: a time which is 2 sd or more above the mean is classified as 'superior'; a. time between 1 sd and 2 sd above the mean is classified as 'above average'; a time within 1 sd of the mean is classified as 'within normal limits'; a time between 1 sd and 2 sd below the mean is classified as 'below average'; and a time 2 sd below the mean is classified as 'impaired'. Data is

provided in the Appendix, Tables A.2 to A.7, for time taken on the Screening Assessment, each of the four ADL tasks and the 'total time taken' to complete SOTOF The data for total time taken represents the sum of time taken on the Screening Assessment and four ADL tasks. The range, mean and standard deviation of time taken (recorded in seconds) are provided in the Appendix, Table A.1. Normative score conversion tables for time provide details of time ranges, time grades (descriptive classifications) and percentage of normal population obtaining time grade.

If you wish to use the normative data to compare your client's performance with a normative sample the client's performance must be timed exactly. Use the tables to provide an indication of the client's speed in relation to the normal elderly population. Results are recorded in the summary sections of the Record Forms.

Descriptive responses

Clients' responses to the items requiring a description may be classified as follows: general descriptions (such as for eating, for washing, to drink with, to wear); concrete examples (such as eating cereal for breakfast, holding, and pouring out milk, keeping warm); imprecise and long-winded descriptions provided by clients with dysphasic problems, such as word-finding difficulties; and responses which appeared to be obviously incorrect. Normative standards for descriptive responses are provided for: items requiring descriptions of the use of objects in the four ADL tasks; items requiring description of the taste of food in the Eating task and drink in the Pouring and Drinking task; and description of the temperature of water in the washing task. Acceptable descriptions are provided (Appendix, Tables A. 8 to A. 14) to aid scoring of these items. With regard to descriptions of the use of objects, each of the ADL tasks involved the use of at least two objects (for example bowl and spoon, cup and jug) and so many descriptions provided by the normative sample contained two pans. Data is, therefore, provided for the first and second descriptions given by normal elderly clients.

Guidelines for observing and scoring the neuropsychological items

Please refer to Further Reading and Other Assessments at the end of this manual for references for other assessment techniques and further literature related to these neuropsychological deficits.

Language

Comprehension: The client responds to auditory stimuli and hearing is intact, however, the client is unable to respond to verbal and/or written commands. The client should be able to use items appropriately but is unable to respond to instructions; problems may present themselves in a similar way to ideomotor apraxia. A less severe comprehension deficit can result in a client being able to respond to simple commands (for example 'What is this?') but not to commands comprising several units/ideas. For example, the command 'Put the cup on the table' contains four units or ideas; the client must understand the meaning of two objects,

'cup' and 'table', an action, 'put', and a position, 'on'. It may be necessary to break down instructions into very simple one-unit commands and then build up the number of units to identify the level of the client's comprehension skills.

Expression: The client is unable to give appropriate verbal responses. This may include problems with naming, sentence construction or difficulties with pronunciation.

Hearing

Hearing acuity: The client is unable to respond to auditory stimuli but is able to comprehend written instructions. Comprehension may be improved by the use of a hearing aid and the tester should sit directly in front of the client and speak slowly and clearly; repetition or rephrasing of commands may be helpful.

Auditory agnosia: The client responds to auditory stimuli (for example will turn head in the direction of a loud, unexpected noise) but is unable to make sense of, and respond appropriately to, any auditory stimuli. Clients with a comprehension deficit may still recognize familiar sounds (such as the noise of a car, an animal or music) and respond accordingly; the client with auditory agnosia is unable to make sense of familiar sounds as well as the spoken word.

Cognition

Orientation: The client is unable to identify self, place and/or time.

Attention: The client is easily distracted from the task. This can present itself in a similar way to short-term memory deficit where the client forgets what she is supposed to be doing; check whether the client can still recall the instruction. Attention span can be affected by tiredness, pain or discomfort; check positioning and give frequent rest periods between tasks and task components.

Short-term memory: This involves the retention of instructions or objects in immediate awareness. The client may hear and comprehend an instruction but be unable to repeat or recall it a few seconds or minutes later. The client may be directed to, see and recognize an object but then forget that it is there; this may be mistaken for, and can be associated with, visual field loss and neglect. The client may forget what she was doing or be unable to remember the second part of an instruction (for example 'Wash your hands in the bowl using the soap, and dry your hands on the towel').

Long-term memory: This represents a permanent record of learned material. The client may not remember that she has already eaten or be unable to recall what was eaten. More severely, the client may understand the concept of dressing and have the skills required but be unable to remember how to dress. The client may be unable to remember a new learnt skill (such as hemiplegic dressing technique) from one day to the next.

Initiation: This involves the ability to start a task through beginning physical or mental activity. Impaired initiation may appear as a slowness of response, decreased productivity, decreased spontaneity or lost initiative. The deficit might be highlighted

by a failure or slowness to respond to commands requiring the initiation of a physical act or a verbal response; for example commands like: 'Eat the food in the bowl using the spoon' and 'Wash your hands in the bowl using the soap, and dry your hands on the towel'.

Motor

Abnormal tone: This can include flaccidity, spasticity, tremor, rigidity and decreased muscle strength. The deficit may be highlighted by a failure to copy action in item 5 on the Screening Assessment or by difficulty with reaching, grasping and manipulating objects in the ADL tasks (for example reaching for spoon, picking up the shirt, lifting cup to mouth or putting down the soap). Problems with grip and coordination may affect the client's ability to identify objects through touch even when sensation remains intact.

Bilateral integration: This involves the interaction of both body sides in a coordinated manner during an activity. The deficit may be highlighted by an inability: to coordinate hands together to lather the soap in the Washing task; to hold the cup whilst pouring liquid from the jug; or to hold a buttonhole open on a garment whilst pushing a button through the opening with the other hand.

Fine motor coordination/dexterity: This involves the use of small muscle groups for controlled movements, particularly in object manipulation. The deficit may be highlighted in any of the test items involving object use, for example use of the spoon.

Sensation

Proprioception: This involves the awareness of the position of body parts and can be associated with sensory discrimination and body scheme deficits. Failure to carry out movement in items ‘Copy what I do’ and ‘Hold this’ on the Screening Assessment, and client's difficulty lifting spoon or cup accurately to her mouth, are examples of tasks which may indicate a deficit. Further clarification is usually needed.

Tactile and temperature discrimination: The deficit may involve light touch, temperature, deep pressure, two-point discrimination and/or pain. Temperature is assessed in Task 2; the client is asked to describe the temperature of the water. Other deficits may be indicated by problems with identifying objects by touch (the tester may need to clarify between tactile agnosia, tactile discrimination and problems with manipulation of the object; further assessment is recommended), problems with maintaining grip or knocking objects. Ask the client to describe textures of the towel, soap and shirt.

Taste discrimination: This involves the recognition and identification of flavours. Taste discrimination is assessed for clients who have expressive language intact through the Eating (Task 1) and Pouring and Drinking (Task 3) 'Can you taste the drink/food?' and 'Describe how it tastes' items. See Normative standards on p.110 (Table A.12) for examples of acceptable descriptions of flavours.

Vision

Visual acuity: The client is unable to identify, through naming or pointing, familiar objects placed directly in front of her. This may occur with visual field deficit and scanning problems. The tester needs to clarify between acuity and object or colour agnosia. Test with large familiar objects held 6 inches away at eye level and then change distance or size of object.

Visual attention: The ability to focus gaze; the client is unable to identify objects visually and does not appear to focus attention on objects. The client may require further testing to differentiate between visual attention and visual acuity.

Visual scanning: The client is unable to scan across the table and, therefore, does not identify all the objects.

Visual field loss: The client is unable to see parts of the table or to see the tester when sitting on the affected side. Further testing may be required in order to define exact deficit (for example homonomous hemianopia, bi-temporal hemianopia).

Visual neglect: The client ignores part of the visual field. For example, the client ignores some items on the table or leaves food on one side of the bowl when she thinks she has finished all the food.

Agnosia

Visual spatial agnosia: May be identified by problems with judging distances and depth or client's inability to orientate herself to the objects on the table.

Visual object agnosia: The client can see the object but is unable to name it or describe its use. The client may be able to describe the object's shape, colour and/or texture. Simultanognosia is a disorder of visual attention that involves impairment in interpreting a visual stimulus as a whole; scenes and objects are perceived in a piecemeal manner. Metamorphosia is a perceptual distortion of objects in which the person reports that linear objects appear curved or discontinuous. The object may be recognised accurately but may appear different (for example in size) than it actually is.

Colour agnosia: The client is unable to name colours, match colours or to point to named colours. Start testing with bright primary colours and then build up to subtle shades and patterns. Remember that colour vision changes with ageing: colour sensitivity shows a gradual loss of fine discrimination brought about by the yellowing of the lens and retinal changes; the blue/green/violet end of the spectrum is particularly affected as the colours are filtered out by the increase in yellow pigments. Visual object agnosia may or may not remain intact.

Tactile agnosia: It is important to differentiate between tactile discrimination, poor motor function (which can restrict the manipulation of objects in the hand sufficient for identification) and tactile agnosia. The client is unable to identify objects through touch although sensation is intact. Ask the client to identify or describe the shape, texture and weight as well as the name of the object.

Apraxia

Constructional apraxia: The client may be unable to put the spoon in the bowl or button into the buttonhole. Clarify with further assessment, for example copying simple drawings, 3D cube designs or laying a tray.

Motor apraxia: The client understands the idea and purpose of the task or movement (for example pick up the spoon, eating) but is unable to carry out the movement although the necessary motor function is intact. The client may be able to carry out simple tasks automatically (for example drink from the cup) but has problems with motor planning and skilled sequence of movement (for example pouring water from the jug into the cup).

Ideomotor apraxia: The client understands the concept of the task or action (i.e. can describe the concept in response to the question 'What do you use these objects for?') but is unable to mime or demonstrate how to use the object. However, the client may be able to carry out the task spontaneously or if the object is handed to her without the command.

Ideational apraxia: The client does not understand the concept, idea or meaning of the task or action and is unable to describe what the objects are used for even when expression is fully intact. Motor performance of tasks may remain intact.

Dressing apraxia: There is some controversy as to whether this is a specific deficit. The deficit usually involves problems with body scheme and spatial relations. The client is unable to organise and put on the shirt, although the necessary motor and sensory function is intact. Where the client has some motor deficit, dressing apraxia is identified when the problems appear to relate more to body scheme and spatial deficits than to motor performance; for example, incorrect organization of garment, putting on the shirt back to front, putting the wrong arm in the sleeve or buttoning the shirt incorrectly.

Body scheme

Somatognosia: The client is unable to recognize and locate body parts; for example, hands for washing, mouth for feeding and drinking, arms and neck for dressing. May occur with spatial relations deficit and lead to dressing apraxia.

Unilateral neglect: The client may only dress one limb, may put both sleeves on one arm, or may only wash one hand. The client may ignore the affected side unless prompted.

Anosognosia: The client appears to ignore or be unaware of any deficit; for example, will attempt tasks which are unrealistic and risk falls or other injury. The client may also deny that the affected limb belongs to her body.

Right/left discrimination: The client is unable to discriminate between right and left, is unable to place objects to the right or left of another object on command and may be unable to identify her own or the tester's right and left body part.

Spatial relations

Figure ground discrimination: The client has problems identifying the foreground from the background; this can result in problems identifying objects,

reaching for objects on the table or finding buttons on a shirt. This deficit may be associated with a poor attention span and the client may be easily distracted.

Position in space: The client is unable to place objects correctly in relation to each other on command or to describe the location of objects in relation to each other or to him / herself. The client may no longer understand concepts of spatial relationships such as up/down, in/out and in front/behind.

Form constancy: The client has problems identifying subtle variations in form and may mistake objects for things with a similar shape, size or colour. For example, the client may identify the jug as a vase or urinal, or the spoon as a trowel or other item of cutlery.

Spatial relations: The client has difficulty perceiving the position of objects in relation to each other and to herself; for example, she may not line up the jug accurately over the cup for pouring, she may not reach accurately for objects or she may be unable to put the spoon in the bowl or her hands into the water.

Depth perception: The client may have difficulty judging the depth of the liquid in the jug or cup and may overfill the cup or may have problems judging the depth of water in the washing bowl.

Distance perception: The client has difficulty judging distances in between objects and between objects and herself. For example, she may reach short of, or beyond, an object or be unable to pour liquid into the cup accurately.

Perseveration

The client repeats actions or sequences unnecessarily. For example, continues to eat or drink after the food or drink has been finished or continues to dry hands when they are already dry.

3 Summary of Development, Standardisation and Technical Data

Summary

The content of SOTOF has been drawn from practice in the fields of occupational therapy and clinical neuropsychology. SOTOF provides a simultaneous assessment of self-care performance and underlying neuropsychological functioning. SOTOF is an observational test that involves watching the client's performance in four self-care areas, simple recording of observations on a two-point categorical scale, and a reasoning component to identify possible underlying neuropsychological deficits. The SOTOF gives reliable results over time, yields consistent results across different clinicians, and relates to other standardised tests.

Content development

The structure and format of the SOTOF have been based on occupational therapy practice and the theoretical framework has been drawn from aspects of occupational therapy, neuropsychological and general systems theories. The SOTOF provides a structured way to administer ADL assessment to evaluate underlying neuropsychological deficit in addition to the client's level of independence in self-care. Unlike many psychological assessments, the SOTOF focuses on more than one construct. The SOTOF represents a standardisation of occupational therapy clinical practice, involves a reasoning component, addresses four levels of function and detects a wide range of dysfunctions. To begin the process of standardisation, and maximize the validity and reliability of the test, procedures for psychological test construction were applied to the development of this test.

Rationale for the development of SOTOF

Both quantitative and qualitative assessment data is needed to form an accurate clinical image of a client, so clinicians often use a range of assessment methods and tests to provide a comprehensive evaluation of the client. In the evaluation of those with neurological damage therapists often administer ADL assessment alongside standardized neuropsychological test batteries. Results from a survey of occupational therapists in practice conducted in 1989 (see Table 3.1) indicated that the majority of occupational therapists valued standardised tests for their quantitative data and information regarding reliability, validity and normative scores. However, the standardised tests available were considered to be based on contrived activities which held little real purpose or relevance for their clients. Occupational therapists reported using many informal observational testing scenarios to collect qualitative information because they felt that tests administered in naturalistic environments increased ecological and face validity. Informal methods were considered to allow

flexibility in the testing procedure and enable the therapist to tailor the context of the assessment to the client's unique requirements. Rigid standardised procedures were not perceived to allow for empathy and exchange. Both this survey of therapists, and an extensive review of literature conducted in 1989 (Table 3.1), established that occupational therapists were using several separate assessment methods. Tests available, at that time, either evaluated neuropsychological dysfunction or occupational performance no tests addressed both simultaneously. The scores obtained from ADL tests related to global functional independence across a range of ADL tasks; none of the tests looked at discrete components of ADL tasks, nor sought to provide explanations for the causes of observed dysfunction. Neuropsychological assessments used by occupational therapists took the format of batteries of discrete sub-tests each addressing specific perceptual and cognitive deficits. None of the neuropsychological tests used by the surveyed occupational therapists were based on structured observations of performance in ADL. To address this gap the SOTOF was developed to provide synchronous evaluation of the client's performance of ADL and the underlying neuropsychological function. The SOTOF was founded on an interdisciplinary conceptual framework drawn from neuropsychological, occupational therapy and general systems theory in order to combine the benefits of psychometric testing with the ecological validity of ADL assessment.

Table 3.1: Summary of research studies to evaluate the psychometric and clinical utility of SOTOF

Study	*Purpose*	*Research questions*	*Design*	*Subjects*	*Results*	*Contribution to instrument development*
Content and construct validation Literature review (1989-90)	Provide construct and content validation for SOTOF based on occupational therapy (OT) and neuropsychology theory, practice and research.	What constructs should SOTOF address? What are accepted operational definitions of these constructs? What items should form the content of SOTOF to adequately address the domain and constructs of interest	Literature review.	Not applicable.	References for literature and published research providing content validation for SOTOF.	First stage of establishment of construct validity: identified and developed operational definitions for constructs to be evaluated. First stage of establishment of content validity.
Content validation Survey (1989-90)	Provide content validation for SOTOF based on occupational therapy practice.	What is current OT use of standardized and non-standardized perceptual and ADL assessment? What are the OT clinician's criteria for a combined ADL-neuropsychological assessment? What will be areas of ADL performance that will comprise the SOTOF test domain?	Structured detailed questionnaire postal survey administered to a sample of occupational therapists with experience in the fields of gerontology and neurology.	Twenty-nine qualified occupational therapists working with stroke patients in England, Ireland, New Zealand and Canada.	OTs were dissatisfied with current perceptual tests; were using informal observation of ADL to screen for perceptual deficits; and felt that the use of ADL had greater ecological and face validity than the use of perceptual tests. A combined ADL-neuropsychological test should be standardized, valid, reliable, quick and easy to administer, inexpensive, portable, relevant to client's sex, age and social status and provide results which relate functional ability to neuropsychological deficit.	Second stage of establishment of content validity: provided criteria which contributed to the test specifications drawn up to direct the development of SOTOF.

Rationale for an observational testing format

The SOTOF is based on an observational testing format. This decision was based on both clinical and theoretical considerations. From a clinical perspective, direct observational testing has been identified as the preferred method, as opposed to self-report and proxy-report methods, for evaluating performance of older clients (Guralnik et al, 1989). Law (1993) expressed concern about the 'significant trend to develop self-report instruments completed by the client rather than evaluations based on direct observation of performance [as] occupational therapists' skill or expertise ... is in assessing clients and drawing inferences based on their direct observation of clients' performance' (p. 234).

This direct observational-based ADL assessment approach was one of the most frequent assessment methods used by occupational therapists to evaluate the function of clients with neurological damage (Ottenbacher, 1980). The functioning of perceptual and cognitive systems cannot be directly observed. Behavioural observation is considered to be an acceptable method for making inferences about these systems. Theory pertaining to the relationship between an individual's observed behaviour and internal organisation also supports the use of observation of performance as an indicator of neuropsychological function. The theoretical foundation of SOTOF, therefore, draws on general systems theory.

The repertoire of a person's acts develops throughout the life span as the result of interaction with the environment. The limitations of such a repertoire means that there are a limited number of responses (output) an individual can make to a given stimulus (input). We learn to associate stimulus and action responses, for example presentation of food with eating. These associations form a conceptual framework of normal stimulus-response interactions. Observation of the demands of an individual's environment provides information about the nature of the information (input) received by that individual. Observation of the output produced in response to this environmental input provides the therapist with an indication of the nature of the client's internal organisation (throughput). The behaviour produced from a well-organized central and peripheral nervous system is relatively predictable, that is, based on an analysis of the situation it is possible to make a reasonable prediction of what a person is going to do before she acts. A perfectly organised system is completely predictable and its behaviour provides no information at all. The more disorganised and unpredictable a system is, the more information you can get by watching it (Miller, 1968). Unpredictable output alerts and orientates the therapist to the possibility of dysfunctional organisation. Unexpected output, in the form of action or language, challenges the therapist and prompts the onset of inquiry: why did the individual behave in this way? Unusual output provides observational cues which prompt hypothesis generation.

The SOTOF is based on an error analysis assessment approach in which the therapist acts as a data processor. The therapist observes the client's behavioural responses to defined stimuli and then selects any unexpected behavioural cues, or observed error, as the focus of the diagnostic reasoning process. Reasons for the observed errors are generated in the form of hypotheses which are then tested against further observational cues, and theoretical and tacit knowledge. For example, if a client is presented with a cup and asked, 'What is the name of this object?' the expected response would be 'cup'. If the client failed to respond, the tester would start to formulate a range of hypotheses which could explain the client's behaviour. Perhaps the client has reduced hearing acuity and did not hear the instruction, or perhaps she heard, but has a language deficit, such as receptive aphasia, and did not understand the instruction. If receptive language and hearing are intact, then the problem might still lie in the language domain, but be one of expressive aphasia. Alternatively, if hearing and language are intact, the problem might have a visual origin, for example visual acuity, visual attention or visual field loss. Further cues to provide information about hearing, vision and language would be sought and then used to evaluate each of these hypotheses. If hearing, vision and language were found to be intact the tester would need to generate further hypotheses to explain the original observed behaviour error (failure to name the cup). A further hypothesis, for example, could be visual object agnosia, which is the failure to recognise familiar objects although vision is intact. All these hypotheses are related to performance component dysfunction. A further explanation could lie with the volitional sub-system.

Dysfunction in a performance sub-system is only one explanation for unexpected output; motivational factors, arising from the volitional sub-system, can also have a profound effect on behaviour. The volitional sub-system should be considered to place observed behavioural cues into the context of the individual's internal, as well as external environment. It is essential to engage the individual's motivation. If judgements made from observational assessment are to be reliable and valid then optimum performance needs to be elicited. The selection of an assessment domain must be made with reference to the interests, roles and habits of the population on which the test is to be used. Motivation may be enhanced by allowing the individual some choice in the assessment activity to be performed. However, the benefits of individual choice have to be balanced against the requirements of standardisation. The SOTOF is based on activities which have some universal human relevance and allows some elements of individual choice (e.g. what they wish to eat and to drink) that do not impact on the reliability of the test.

Effective use of observational assessment is highly dependent on the quality of the therapist's reasoning. One of the aims for the development of SOTOF was to provide a structured format to guide the therapist's reasoning during the observational assessment.

The SOTOF diagnostic reasoning process

Occupational therapists use a diagnostic reasoning process to identify neuropsychological deficits during the observation of the ADL tasks. Diagnostic reasoning is a component of clinical reasoning and involves the creation of a clinical image of the client through cue acquisition, hypothesis generation, cue interpretation and hypothesis evaluation (Rogers and Holm, 1991). The process provides a summary of a client's deficits in terms of occupational performance and performance components. The four stages of the process are outlined below with examples (given in parentheses) drawn from SOTOF.

1. A description of the problem (client unable to feed independently).
2. Postulates possible causes for deficits (motor deficit such as hemiplegia or perceptual deficits such as visual object agnosia and altered body scheme).
3. Identifies the cues which lead to recognition of the problem (client unable to name or describe use of objects, client unable to reach for and grasp spoon, client unable to indicate position of mouth).
4. Renames the underlying pathology (stroke or dementia).

SOTOF provides a structured method for organising and guiding thinking during each of the diagnostic reasoning processes. Within this structured format, a balance between standardised procedure and flexibility has been achieved. Sufficient flexibility is allowed to enable the therapist to tailor the assessment to the needs and performance of each individual client. To guide the therapist's diagnostic reasoning SOTOF contains:

- a clearly defined data field, which is provided through a description of ADL assessment domains;
- a breakdown of the ADL domains into discrete behavioural components, which are described in each test item on the Instruction Card (see example in Figure 2.1) and are listed in the Record Forms;
- a structured Instruction Card for systematically searching the data field for cues. This involves searching one discrete behavioural component at a time in the order in which these components are sequenced in the normal performance of the ADL task (see Figure 2.1 for example Instruction Card);
- a list of possible deficits which could account for dysfunctional performance in each of the discrete behavioural components, which are given for each test item in the middle column on the Instruction Card;
- operational definitions for all the neuropsychological deficits which may be observed through the performance of the ADL task; these are provided in the Glossary (pp.129-133);
- examples which provide a knowledge base from which the therapist can draw information to aid hypothesis generation and evaluation; these are provided through scoring guidelines (pp. 35-40);

- a system, comprising a Record Form, for categorising cues into skills and deficits, for noting hypotheses and for recording further cues acquired when there is a differential diagnosis (see Figure 2.2 for example Record Form);
- a system, comprising a checklist of all neuropsychological deficits which may be observed from the performance of the ADL domain, for recording diagnoses (see Figure 2.3 for part of the Neuropsychological Checklist);
- information on other assessment tools which could be used alongside SOTOF to collect additional cues when a differential diagnosis remains at the end of testing (see Further Reading and Other Assessments which can be found at the end of this manual on pages 133-134).

Development of the ADL domains covered by SOTOF

ADL is an enormous domain of human activity and was carefully reviewed and defined for the test. Even a single area of ADL, such as dressing, can be viewed as a whole comprising many sub-parts. Dressing includes the tasks of deciding what to wear, collecting items of clothes from their storage places, correctly sequencing the order in which clothes are donned, and putting on each individual item of clothing. Studies have found that therapists often select one indicator of function rather than assessing a whole ADL domain (for example Rogers and Masagatani, 1982). Initial research had shown that therapists usually assess more than one area of ADL, such as washing and dressing. In the frequent case of differential diagnosis, provisional hypotheses can be further evaluated during the observation of a second task through the collection of specific additional cues. SOTOF is therefore based on four ADL tasks. ADL assessment usually begins with the observation of the domain of personal ADL. Five personal ADL sub-domains emerged frequently in the literature and survey responses (see Table 3.1): these were dressing, washing, feeding, drinking, and grooming. The sub-tasks included in each of these sub-domains were listed. For example, feeding may involve feeding with hands (eating a sandwich), feeding from a bowl using a spoon, spatula or chopsticks (eating rice dishes, soup, cereal and puddings), and eating from a plate using a knife and fork (eating slices or joints of meat). From the list of the sub-tasks of these five ADL sub-domains the four SOTOF ADL tasks were selected based on the following criteria.

1. As SOTOF is a standardisation of current occupational therapy practice, the tasks selected are commonly used by occupational therapists for the assessment and treatment of neuropsychological deficit. These tasks are familiar to clinicians, and this aids the recognition and analysis of unexpected, dysfunctional performance.
2. To ensure ecological and face validity for a wide range of clients, the tasks selected are functional, purposeful, familiar and have relevance for both sexes and for people from different ethnic and social backgrounds; the tasks are also suitable for administration in the client's own environmental context.

3. SOTOF addresses the specific age and requirements of the client population for which it was developed; the tasks are, therefore, appropriate for older people with neurological diagnoses (e.g. stroke); tasks are age-appropriate, provide a suitable level of challenge and can be performed one-handed and from a supported seated position.
4. Short test administration and cost had been common features of many therapists' criteria for clinically useful tests; the tasks, therefore, take no longer than ten minutes for a healthy person to perform and involve the use of common, easily obtained, inexpensive objects and materials.

The following four tasks were selected to form the basis of SOTOF:

1. Eating from a bowl using a spoon.
2. Washing hands in a bowl of water.
3. Pouring liquid from a jug into a cup and drinking from the cup.
4. Putting on a long-sleeved, front-fastening upper garment.

Therapists often conduct ADL evaluation over several assessment sessions owing to time constraints and because older clients often tire quickly. This practice was taken into consideration when deciding on the structure of SOTOF. Each of the four ADL tasks were used as the basis for a discrete subtest with its own Instruction Card and Record Form.

Development of the constructs addressed by SOTOF

SOTOF addresses four broad constructs: perception; cognition; motor; and sensory function. Constructs are identified through behavioural cues (skills and abilities) which are gathered from the observation of occupational performance in the domain of the four discrete personal ADL tasks. The specific content of the test comprises neuropsychological constructs and the behaviours representing those constructs. The behaviours normally expected to occur during functional performance of each of the selected ADL tasks were identified. These behaviours were linked to the constructs which SOTOF addresses. Two methods were used to identify the content (behaviours and constructs) of the test: a review of occupational therapy, medical and neuropsychological literature; and activity analysis. The constructs addressed by SOTOF were identified through a review of the research literature regarding deficits arising from neurological damage, and through critiques of current occupational therapy neuropsychological assessments.

The literature review and test critiques revealed four main performance components commonly addressed by occupational therapy neuropsychological assessment: these were perception, cognition, sensation and motor function. Review of literature and assessments provided information on the specific neuropsychological deficits theoretically linked to each of these four constructs. A list of deficits which SOTOF addresses is given below. Perceptual deficits have been divided into four main

groups: agnosias; apraxias; body scheme deficits; and spatial relations deficits (Zoltan et al, 1986). Perceptual deficits are listed under each of these four headings.

1. Agnosia - visual spatial agnosia, visual object agnosia, colour agnosia, tactile agnosia,
2. Apraxia - constructional apraxia, dressing apraxia, motor apraxia, ideomotor apraxia, ideational apraxia.
3. Body Scheme - somatognosia, unilateral neglect, anosognosia, right/left discrimination.
4. Spatial Relations - figure ground discrimination, position in space, form constancy, spatial relations, depth and distance perception.

Other deficits relating to the cognitive, sensory and motor performance components are:

- Cognitive deficits - orientation, initiation, attention and memory.
- Language deficits - receptive and expressive aphasia/dysphasia.
- Auditory deficits - hearing acuity and auditory agnosia.
- Visual deficits - visual acuity, visual attention, scanning, visual field loss, visual neglect.
- Other sensory deficits - proprioception, tactile, temperature and taste discrimination.
- Motor deficits - abnormal tone, bilateral integration and fine motor coordination/dexterity.
- Perseveration.

Operational definitions for all these neuropsychological deficits are provided in the Glossary (see pages 129-133). This glossary has been reviewed and updated for the 2nd edition of SOTOF

The behaviours that represent the constructs evaluated by SOTOF

Observation is a method used for identifying test behaviours (Crocker and Algina, 1986). To identify the behaviours that represented the constructs evaluated by SOTOF, all four of the ADL tasks were subjected to a detailed activity analysis. The identification of discrete behaviours was undertaken to increase the sensitivity of the test. Many ADL assessments simply evaluate whether a client is independent or dependent in a range of ADL tasks. These instruments answer questions concerning 'what' an individual can do, and do not provide data regarding the underlying causes of dysfunctional performance. Scales defined by small increments are considered to be more sensitive (Fisher, 1992). The global ADL tasks need to be sub-divided into measurable components to increase sensitivity. Activity analysis identified discrete behavioural components for the SOTOF tasks. This process identified the sequence of actions, the skills required (observed behaviour) and the required underlying neuropsychological functioning needed to perform these skills (constructs).

Two strategies were used to increase the objectivity of the analysis. First, a comprehensive, structured activity analysis protocol was selected (Pedretti, 1985). Secondly this Instruction Card was used as a basis for detailed activity analyses which were conducted by a sample of six occupational therapists. The Instruction Card selected covered all constructs of interest, that is, the sensory, perceptual, cognitive and motor demands of the activity. It also provided a description of the component steps of the activity, considered safety factors, sociocultural symbolism, psychological-emotional responses to the activity, and was structured in a question format that was easy for therapists in the study to follow. The data from the analyses was used to identify component steps, input received from the materials and performance of the task, the expected behavioural responses to each of the tasks, and the nature of any feedback. The following information was listed for each of the tasks.

- Description of component steps of the task - including the normal sequence of steps and any repetition of sequences of steps.
- Description of motor responses - specific descriptions of required movements including muscle groups and joints involved, range of movement and degree of strength required, and repetition of specific movement patterns.
- Description of the sensory input from the task materials and performance of the task - including tactile, proprioceptive, vestibular, visual, olfactory, gustatory, pain, pressure and thermal stimuli.
- Description of required tactile-proprioceptive-vestibular functioning including equilibrium and protection reactions, postural and bilateral integration, tactile discrimination and proprioceptive feedback.
- Description of the visual functions required - including visual scanning, and recognition and differentiation of colour, size, shape and form.
- Description of requirements for perception of spatial relations - including recognizing, differentiating, matching and fitting shapes, forms and patterns.
- Description of requirements for figure-ground discrimination.
- Description of requirements for gross and fine visual-motor coordination.
- Description of the auditory functions required - including hearing acuity and sound differentiation.
- Description of the cognitive demands of the task - including the need for long- and short-term memory, sequencing, problem-solving ability, concentration, generalization of learning and comprehension.

The SOTOF scoring format

The SOTOF has a numerical scoring system linked to the 6 level graduated mediation protocol. Many assessment forms allocate numerical labels to nominal or ordinal scores. It should be remembered that with an ordinal scale the distance between items on the scale is unknown. Some scales have an aggregate score, based on the sum of numerical scores which are essentially derived from ordinal

data. It should be noted that two persons could receive the same total score although their functional profiles, which have important implications for occupational therapy, are quite different, so therapists should not look at the summed scores alone.

The purpose of SOTOF was to provide data on several levels (overall ability to perform task, ability to perform specific skills and neuropsychological function). The main scoring system used to record observations from the Screening Assessment and four ADL tasks is a categorical scoring system.

Categorical scoring system for the standardised phase

The occupational therapy process focuses on skills as well as deficits, thus 3 the results obtained by SOTOF reflect both what the client can and cannot do. Each ADL task has been submitted to a detailed activity analysis to identify its discrete behavioural components. A nominal, two category scoring system is used for the task components. This is a simple scoring system, based on dichotomous decisions, which are recorded easily by ticking boxes on the Record Forms. Measurement of a client on SOTOF is interpreted in terms of defined component behaviours, which the client either does or does not exhibit. Each of the recorded behaviours forms a criteria that is reviewed in terms of a dichotomous decision. SOTOF has two decision categories, 'able/yes' and 'unable/no'. This dichotomous categorical format provides a picture of both the client's abilities and deficits. Scores are recorded on the Screening Assessment record form (if the screen is used) and the record form for the four ADL tasks. The SOTOF record forms detail each behavioural component alongside boxes for each scoring category. The tester simply ticks the appropriate box to indicate whether the client was able or unable to perform the task component.

Test specifications

SOTOF is based on the following test specifications.

- Each test item involves a discrete behavioural component.
- The instructions to elicit each behavioural component are clearly outlined in the Instruction Card.
- Each behavioural component is recorded using a dichotomous 'able' or 'unable' categorical scoring system.
- In addition to the categorical scores, space is given alongside each behavioural component to record qualitative data on the performance of the component, to note hypotheses and to identify further testing required to collect any additional cues needed for the evaluation of the hypotheses.
- The most common neuropsychological deficits associated with the failure to perform a behavioural component are identified on the Instruction Card as an aid to initial hypothesis generation.

- Advice for the collection of additional observational cues to assist the evaluation of hypotheses, in the case of differential diagnosis, is given on the Instruction Card.
- Space is provided to record qualitative information, such as data on the client's ability to perform each task and any relevant factors which could affect performance (such as a noisy test environment).
- A glossary of possible neuropsychological deficits is provided as an aid to hypothesis generation and evaluation.
- A checklist of possible neuropsychological deficits is provided to record diagnoses at the end of testing.
- Each neurological deficit identified through the reasoning process is recorded using a dichotomous 'present' or 'absent' categorical scoring system.

Standardisation

SOTOF can be used as both a criterion-referenced and a norm-referenced test, which is comparing the client's performance with the performance of a clinical group or with the normal population, respectively. Norm-referenced tests are usually employed when the limits of achievement are not clearly defined and an individual's ability may progress almost without limits, in functions such as critical thinking and originality for example (Anastasi, 1988). Concepts of independence and dependence can be clearly defined and observed. Criterion-referenced mastery-based testing is applicable when evaluating independence in basic skills, such as those selected for SOTOF from the domain of personal ADL. This type of ADL test is rarely norm-referenced for adults because therapists assume that normal performance in ADL is the ability to do activities completely (Law, 1993). This is a dangerous assumption to make when assessing older clients as the functioning of all performance component systems alters with ageing. As ageing impacts on occupational performance, normative data for performance on SOTOF was required. Performance is often evaluated in terms of both the ability to complete a task and the speed of performance. As the limits, in terms of speed of a client's performance, are not clearly defined, and speed of performance is affected by primary ageing, it was important that time data should be norm-referenced. A normative control study was undertaken to investigate how individuals who do not have neurological disorders perform on SOTOF (Table 3.2). Normative standards for time data, and expected responses to items requiring a description from the client, are provided in the Appendix.

Validity, reliability and clinical utility studies

Several studies have been undertaken to examine aspects of the validity, reliability and clinical utility of SOTOF These include content validity, construct validity, criterion-related validity, face validity, internal consistency, interrater reliability, test-retest reliability, and clinical utility.

Content, construct, criterion-related and face validity: Therapists need to know whether a test adequately represents the performance domains and/or constructs they are interested in, and whether the test items address these domains and/or constructs in the correct proportions. Content validity refers to the degree to which a test measures what it is supposed to measure judged on the appropriateness of its content (Bartram, 1990). Construct validity involves the extent to which a test can be said to measure a theoretical construct or constructs (Anastasi, 1988). Criterion-related validity relates to the effectiveness of a test in predicting an individual's performance in specified activities. This is measured by comparing performance on the test with a criterion that is a direct and independent measure of what the test is designed to predict (Anastasi, 1988). Both predictive and concurrent measures can be used to determine criterion-related validity. In the case of SOTOF criteria selected related to ADL function and neuropsychological functioning as measured by existing, established ADL and neuropsychological tests (Table 3.3). To engage the motivation of the client to carry out a test to the best of her ability, it is important that the test appears relevant and acceptable to her. Face validity concerns the acceptability of a test to the test-taker and the degree to which a test-taker sees the test as reasonable and appropriate (Bartram, 1990). Several studies were undertaken to explore these types of validity. The purpose, research questions, design, subjects, sample sizes and results of these studies are provided in Tables 2.3 (page 25), 3.1 (page 43), 3.2 (page 54), 3.4 (page 56) and 3.5 (page 58).

Table 3.2: Summary of research studies to evaluate SOTOF's psychometric properties & clinical utility

Study	Purpose	Research questions	Design	Subjects	Results	Contribution to instrument
Interrater and test-retest reliability (1991-93). Funded by St George's Hospital Special Trustees: April to September 1991 and by a Nuffield Foundation Major Research Grant; January 1992-93	To establish the interrater and test-retest reliability of SOTOF.	Is there correlation between: scores obtained by two different raters scoring the same subject test administration; and scores obtained by one rater administering the test to one client on two separate occasions?	First day two OT raters independently score a client's performance on SOTOF, test is administered by one of the raters. Second day the rater, who administered test on first day, retests the patient under same test conditions. Results compared to reliability values quoted in other test manuals and research articles.	Thirty-two OTs and 37 patients in the UK (diagnoses = 21 stroke, 15 dementia, and 1 head injury) of which 54.1% were female and 45.9% male, age range 60-91, mean 75.6 years.	Statistical analysis included computation of percentage agreement, chi-square and Kappa. Screening Assessment has high test-retest (97.7%, 0.92) and interrater (97.5%, 0.94) agreement. The average values for the four ADL tasks range 90.3-93.8%, 0.5-0.77 for test-retest and 89.5-91.6%, 0.37-0.67 for interrater reliability. The Neuropsychological Checklist has acceptable levels of reliability: 95.2%, 0.55 for test-retest and 95.2%, 0.54 for interrater reliability.	Established acceptable levels of test-retest and interrater reliability for Screening Assessment, four ADL tasks and Neuropsychological Checklist as compared to other standardized ADL and neuropsychological tests.
Normative study (1992-93). Funded by Nuffield Foundation Major Research Grant: January 1992-93.	Obtain normative standards on the performance of a representative sample of clinically healthy older adults. Performance to be evaluated in terms of ability to fulfil each task requirement and the time taken to perform each task. Compare, data to that of patient samples.	Are all clinically healthy older clients able to perform all the SOTOF test items? What percentage of the normative sample exhibits each of the neuropsychological deficits? What is the range of times taken by the normative sample to perform the SOTOF tasks?	SOTOF was administered, by five trained OT research assistants, to a sample of clinically healthy older clients. Data on time taken for the Screening Assessment and four ADL tasks was recorded, along with full notation of subject's responses to descriptive items.	Eighty-six Caucasian adults (68.6% female, 31.4% male) with no history of neurological deficits, drawn from south-east England. Age range 60-97, mean 73.5 years. Representative for socioeconomic status.	Between 1.2% and 10.5% of the normal sample failed 29% of the 100 SOTOF ADL task items and 33.3% of the neuropsychological deficits were highlighted (the majority of these related to visual/ auditory acuity, and colour/ tactile discrimination as expected with normal ageing). Normative time data and verbal responses for descriptive items were obtained.	Established normative standards for performance on SOTOF (time and ability). SOTOF found to discriminate between patients with neurological damage and healthy adults for both ability and time taken. Established normative standards for descriptive responses.

Table 3.3: Key to concurrent and predictive measures

Measure	Full test name, authors and dates
RPAB	Rivermead Perceptual Assessment Battery (Whiting et al, 1986; Lincoln and Edmans, 1989).
NART	National Audit Reading Test (Nelson, 1982; Nelson and Willison, 1991).
COTNAB	Chessington OT Neurological Assessment Battery (Tyerman et al, 1986; Laver and Huchison, 1993).
MEAMS	Middlesex Elderly Assessment of Mental State (Golding, 1989).
RADL	Rivermead ADL Assessment for Stroke Patients (Whiting and Lincoln, 1980).

Internal consistency: Therapists often need to generalize from the performance of a subject on test items to the broader domain of behaviour from which these items were drawn. Procedures to evaluate this type of reliability are called internal consistency methods and a study of this nature was undertaken to evaluate the internal consistency of the SOTOF The SOTOF is based on four sub-tasks of personal ADL, and it is reasonable for therapists to want to generalize from the performance of these sub-tasks to the broader domain of personal ADL. Internal consistency was evaluated by matching SOTOF items that were designed to test the same or similar performance domain or construct in each of the four ADL tasks and comparing clients' performance on these matched items. Information summarizing the purpose, research question, design, subjects, sample size and results from this study is provided in Table 3.5.

Test-retest and interrater reliability: Therapists need to know whether a test is consistent across raters and time. If a test is to be used for evaluative purposes it must have high test-retest reliability Both interrater and test-retest reliability are frequently addressed in test manuals and articles describing new tests (Ottenbacher and Tomcheck, 1993). Studies were, therefore, undertaken to evaluate both these types of reliability. Test-retest reliability was evaluated with pairs of therapists and clients; the SOTOF was administered to the same client by the same tester on two separate testing occasions held one day apart. Interrater reliability was evaluated with trios comprising a pair of therapists and a client; they independently scored the same test administration of the SOTOF to a client, one therapist administered and scored the test and the other observed the test administration and scored the test. Information summarizing the purpose, research questions, design, subjects, sample sizes and results from these studies is provided in Table 3.2.

Table 3.4: Summary of research studies to evaluate SOTOF's psychometric properties and clinical utility

Study	Purpose	Research questions	Design	Subjects	Results	Contribution to instrument
Construct and criterion related validity (1991-93). Funded by Nuffield Foundation Major Research Grant: January 1992-93.	To establish construct validity through comparison of SOTOF with other recognized measures of the same theoretical constructs. To establish criterion-related validity through comparison of SOTOF with criteria obtained using concurrent and predictive measures.	Is there a relationship between the subjects' performance on the SOTOF and: (1) other measures of the ADL domains; and (2) other measures of neuropsychological constructs?	Subjects administered the following measures in random order: SOTOF; RPAB, shortened version; NART; COTNAB, 3 items; MEAMS; and RADL for stroke or an ADL interview (developed for the study) administered to patients' nurse or OT,	Twenty-two patients (63.6% female, 36.4% male) with a primary diagnosis of stroke. Age range 62-92, mean 76.2 years. Subjects were, in-patients in hospitals in London and Kent.	Statistical analysis involved the computation of frequency of deficits across all measures, and Chi-square, Fisher's exact test and Phi coefficient for matched items from SOTOF and each of the concurrent and predictive measures. Overall dysfunction in ADL 'performance and identification of neuropsychological deficits on the SOTOF was mirrored by the identification of dysfunction and deficits on all the other measures. SOTOF appears to relate highly (<0.05 level) to measures of ADL and to some items from the neuropsychological tests.	Established construct and criterion-related validity.

Clinical utility

Even though a test might be thoroughly standardized, valid and reliable, it will be useless if it does not have application to the clinical setting for which it was designed. Therapists are concerned about the practical application of a test, including how long it takes to administer, how heavy it is to carry and the level of expertise required to administer it. Clinical utility studies encompass issues of instructions, cost, time, acceptability and format (Law and Letts, 1989). Cost relates to financial outlay, and to the therapist's and client's time and energy expenditure. These costs are reflected in the amount of equipment, space, training time and expertise necessary to administer and interpret the results of an assessment. The format and clarity of the Instruction Cards and Record Forms also have a direct bearing on the ease with which a test is understood and administered. In addition, the test should be acceptable to the therapist, as well as the client, who should understand and agree with the usefulness of the items being measured. All these issues were addressed in a clinical utility study in which 44 occupational therapists administered the SOTOF to at least one of their stroke patients and reviewed the ease with which the test was administered, the time taken, the acceptability of the materials and the items, and the relevance of the test to their clinical setting and client group. Information summarizing the purpose, research question, design, subjects, sample size and results from this study is provided in Table 3.5.

A more detailed description of the findings of the Test-rest and inter-rater reliability studies for the SOTOF (1st edition) can be found in Chapter 4.

A more detailed description of the findings of the Clinical utility and Face validity studies for the SOTOF (1st edition) can be found in Chapter 5.

Table 3.5: Summary of research studies to evaluate the psychometric properties and clinical utility of SOTOF

Study	Purpose	Research questions	Design	Subjects	Results	Contribution to instrument
Face validity Client survey (1991). Funded by South West Thames LORS grant and St George's Hospital Special Trustees	To evaluate the degree to which the client perceives the SOTOF as being a reasonable and appropriate test of their functioning.	What are clients' perceptions of the purpose of SOTOF? Are these tasks perceived as usual, familiar activities? What are the clients' experiences of taking the test?	SOTOF manual and feedback structured interview provided in a postal survey to OTs who administered the test and the interview to a client with stroke.	Forty-four clients with a primary diagnosis of stroke being treated in hospitals throughout the UK and Ireland.	95% of clients felt the SOTOF ADL tasks represented activities they would normally do and none of the clients minded being asked to do the tasks. The majority found SOTOF to be interesting (87.5%), useful (87.5%) and enjoyable (85%). Only 12.5% found the test at all stressful.	Establishment of face validity.
Clinical utility Occupational therapist survey (1991). Funded by South West Thames LORS Grant.	To evaluate the clinical utility of SOTOF, in terms of instructions, cost, time, acceptability and format, from the perspective of the therapists.	Research questions pertained to the utility related to the test manual and materials, administration time, relevance to client, test-induced anxiety and level of expertise required to give test.	SOTOF manual and feedback structured questionnaire provided in a postal survey to peer OTs who administered and reviewed the test.	Forty-four qualified OTs and 48 stroke patients located across the UK, Ireland and Belgium.	SOTOF was: easily understood and administered; relatively quick to administer as compared to current standardized batteries; relevant for clients; not unduly stressful for clients; and suitable for all qualified occupational therapists.	Establishment of clinical utility.
Internal consistency Data obtained through occupational therapist survey (1991). Funded by South West Thames LORS Grant.	To evaluate whether performance on the test can be generalized to the whole ADL and neuropsychological domain, by examining the consistency of performance on comparable test items	Is there consistency between subjects' performance on similar (matched) items from the four ADL tasks and deficit items for each ADL task on the neuropsychological checklist?	SOTOF was administered to stroke patients by qualified OTs in hospitals in UK and Ireland. Completed test forms were returned for analysis.	Thirty-seven clients with a primary diagnosis of stroke being treated in hospitals in UK and Ireland.	Statistical analysis involved computation of Fisher's exact probability test. The majority of matched items from the Neuropsychological Checklist and some matched items from the four ADL tasks and screen were significantly related at the <0.05 level for two-sided probability.	Established internal consistency across whole test as shown by high levels of consistency on the Neuropsychological Checklist. Variability across four ADL tasks indicates that all tasks should be given.

4 The Test-retest and Inter-rater Reliability of SOTOF (1st edition)

Summary

The purpose of this study was to evaluate the test-retest and inter-rater reliability of the original version of the SOTOF. The method involved the examination of the correlation between (1) scores obtained by pairs of occupational therapist raters scoring the same administration of the SOTOF to an older person (research participant); and (2) scores obtained by one occupational therapist rater administering the SOTOF to the same person on two separate occasions one day apart. The sample comprised of 32 occupational therapists and 37 older people. The sample comprised 54.1 percent females and 42.9 percent males, aged between 60 and 91 years. The majority (n = 21) of these patients had a primary diagnosis of stroke, 15 had dementia and 1 had a head injury. Several statistical analyses were undertaken; these included Percentage agreement, Pearson's Chi-square, Fisher's exact test, Phi Coefficient and Cohen's Kappa. Results indicated that both the average percentage agreement and approximate average Kappa values obtained on the SOTOF's sub-tests and Neuropsychological Checklist compared favourably with other Occupational Therapy standardised assessments. The SOTOF Screening Assessment appeared to have very good test-retest reliability (97.7 percent, Kappa approximate value of 0.92) and inter-rater reliability (97.5 percent, Kappa approximate value 0.94), and can be used as a reliable indication of gross motor, visual and cognitive functioning. The four SOTOF ADL Tasks have higher inter-rater reliability (90.3-93.8 percent, Kappa: 0.5-0.77) than test-rest reliability (89.5-91.6 percent, Kappa: 0.37-0.67). Examination of the reliability of the Neuropsychological Checklist found that the average percent agreement for test-retest reliability was 95.2 percent (approximate average Kappa value was 0.55) and inter-rater reliability was very similar at 95.2 percent (Kappa 0.54).

Types of reliability

This study focused on the evaluation of two types of reliability; test-retest reliability/consistency and inter-rater reliability/agreement. Reliability has been defined as the "consistency or stability of empirical indicators between raters or from one measurement to another ...it is the extent to which a measurement is free from random errors, ...it can be broadly defined as the consistency of a measurement" (Ottenbacher and Tomchek, 1993, p. 10). **Inter-rater reliability/agreement** refers to the "agreement between or among raters" (Ottenbacher and Tomchek, 1993, p. 11). Patients might be referred from one setting to another (e.g. ward to day hospital), or be re-referred after discharge. This can result in the need for a patient to be assessed by several different occupational therapists over a period of time. When this occurs it is important to gauge how likely a change in a patient's performance on a test is a result of a change in rater as opposed to a genuine change in the patient's level of ability. **Test-retest reliability** has been defined as the "correlation between the scores obtained by the same person on the two administrations of the test" (Anastasi, 1988, p. 116), and as the "consistency of an evaluation or test score over

time" (Ottenbacher and Tomchek, 1993, p. 11). A similar methodology is used to evaluate both test-retest reliability and intra-rater reliability. Intra-rater reliability/agreement refers to "the consistency of judgements made by the same rater over a period of time" (Ottenbacher and Tomchek, 1993, p. 11). Frequently, an occupational therapist will wish to evaluate the effectiveness of a treatment programme by re-testing a patient on an assessment administered prior to treatment to see whether desired changes in function have occurred. It is, therefore, important that changes in a patient's performance on the test are not affected by the time interval or by the rater. A study was conducted to provide a measure of both the inter-rater and the test-retest / intra-rater reliability of the SOTOF.

Methods for evaluating reliability

Measurement of a subject on the SOTOF is interpreted in terms of the defined criterion behaviours which the person may or may not exhibit. If a subject is able to perform, and therefore pass, all the items in a task then that subject is considered to be independent for that task. The individual is not considered to have underlying neuropsychological deficits in any of the performance components which would impede his or her occupational performance in the Task's ADL domain. Criterion assessments usually have one of two main purposes: estimation of the domain score, i.e., the proportion of items in the domain which the subject can pass correctly; or mastery allocation. In mastery allocation the domain score is divided into a number of mutually exclusive mastery categories which are defined by cut scores. The observed test results are used to classify subjects into the mastery categories. "The most commonly cited example has one cut score and two categories, master and non-master" (Crocker and Algina, 1986). The concept of mastery allocation to one of two categories is applied to all the test items. The first phase standardized element of SOTOF uses a dichotomous, nominal scoring system; for each item there is an understanding of what the subject should be able to do in order to be classified in the master category which is labelled as 'able', conversely, failure to perform the item to this specified level results in the classification of non-master or 'unable'. The data produced from each SOTOF item is therefore categorical and based on the judgement made by the therapist regarding the subject's ability or inability to perform the item. The evaluation of the reliability is concerned with the consistency or accuracy of the classification decisions made from the observation of the subject's performance. Analysis requires the application of a statistic to a two by two contingency table constructed for each item for (1) the first and second administration carried out by the same rater and (2) the same test administration scored by two different raters.

Reliability study Research Questions

1. Is there correlation between scores obtained by two different occupational therapist raters scoring the same administration of the SOTOF to one patient? This question focused on the inter-rater reliability of the SOTOF.
2. Is there correlation between the scores obtained by one occupational therapist rater administering the SOTOF to the same patient on two separate occasions one day apart? This question sought to establish the test-retest and intra-rater reliability of the SOTOF.

Methodology

The evaluation of reliability involved a combined sample obtained from two separate studies using the same methodology.

First Reliability Study: Identifying and sampling the population

The sample population was drawn from two groups: qualified, hospital based occupational therapists working with older people with a diagnosis of stroke; and patients aged 60 years and over with a primary diagnosis of stroke. Patients with a recent onset of stroke are one of the target populations for the test. For this study, testing was to be undertaken no more than 12 weeks (3 months) from the onset of the stroke. The participants were drawn from a sample of occupational therapists recruited to the research as a result of a letter published in the British Journal of Occupational Therapy. The occupational therapists were contacted by telephone to take part in the reliability study. They were asked if they had a colleague who would be able to carry out the research with them. The therapists were working in hospitals within the United Kingdom.

Procedure

Ethical approval was provided by the St George's Hospital Medical School Ethics committee. Therapists agreeing to assist with the study were sent a packet comprising a letter, questionnaire, test manual and three sets of assessment forms. The letter gave details of the purpose of the study and the procedure. The questionnaire covered: (1) occupational therapists' details including year qualified, current clinical area, grade, experience working with elderly and stroke patients; (2) therapists' prior knowledge of the patient with an outline of previous intervention; (3) patients' details including their age, sex, primary / secondary diagnoses, and date of onset of stroke; and (4) assessment details including date, time and location of testing. The study was undertaken over a two-day period. On the first day, one of the occupational therapists administered the assessment to the patient and the second therapist observed the test administration. The two therapists were instructed to independently record their observations on the SOTOF Observational Task checklists and the Neuropsychological checklist. It was essential that there was no collaboration or conferring between the therapists. On the second day, therapists were instructed to have both tests administered by the same therapist, in the same test location and at the same time of day. They were told to record the testers' initials, date, time and location of testing for both test administrations on the questionnaire. The first author was available for clarification.

Other test developers have used video tapes of patients taking a test, completed test forms or drawings, and photographs of different arrangements of test items, to measure inter-rater reliability. These tapes, forms or photographs are scored by a number of different raters (e.g. Whiting et al, 1985). As the SOTOF involves the observation of four complete tasks, as well as the Screening assessment items, it would be difficult for a rater to gain a complete picture of the subject's performance from one frame or angle. It was impractical to film and edit videotape that had been shot from several angles. The participant does not complete any written or drawn items on the SOTOF and as the test involves the observation and evaluation of a

person's action rather than an end product, (such as a those produced with block design or card sequencing items), photographing test items was also inappropriate. The SOTOF involves on-going clinical reasoning during the assessment. For example: decisions regarding the need for prompts or cues, such as the action 'on command' or 'when handed' object items; or the evaluation of language with the subsequent selection of different administration methods for some items dependent on whether the person has expressive language intact, such as the colour and object recognition items. Because of the nature of the test it was decided that people with varying levels of function should be tested and that the actual administration of the test should be observed by a second therapist. The two therapists (raters) agreed not to confer. However, it should be noted that the clinical reasoning element of the SOTOF is such that the observer could form opinion concerning the patient's function from the way the therapist gives certain test items. For example, if the therapist asks the patient to identify items though pointing rather than naming, the observer could determine that the patient has problems with expressive language.

Second Reliability Study: Identifying and sampling the population

Additional participants were recruited from two hospitals in the south-east of England. Canterbury and Thanet Health Authority Ethical Committee approved the collection of data on the SOTOF for reliability, concurrent validity and normative studies, with participants who were clinically healthy people and/or had primary diagnoses of stroke, dementia, head injury or Parkinson's disease. The diagnostic categories for patient samples were increased at the request of occupational therapists that had taken part in earlier studies and felt that the SOTOF had relevance for an expanded population. Both in-patients and day-patients, under the care of local geriatricians and psychogeriatricians, were recruited for this study. One full time and three part-time occupational therapy research assistants were employed. Participants were recruited through referral from local consultant geriatricians and occupational therapists. The research assistants attended ward rounds and meetings in order to identify suitable patients for the study.

Procedure

Once identified, the researcher visited potential participants on the ward or day hospital and provided an information leaflet outlining the project. A verbal explanation of the nature and purpose of the study was also provided at this stage. Potential participants were given time to discuss the project with their carers, relatives and/or friends and to read the information. When potential participants had visual or language deficits the leaflet was read out loud to them by the researcher or a member of their multidisciplinary team. Patients with stroke were to be tested on the wards and in the occupational therapy department of a local hospital; patients with dementia were to be tested at a psychogeriatric day hospital, on the wards of a second local hospital or at their own home. Prior to testing, the participant signed two copies of the consent form; one copy was attached to the patient's medical notes and the other was attached to their research records. The same testing procedure followed for the first reliability study was used for this study to allow the valid combination of the two samples for the statistical analysis.

Description of sample and testing situation for the first study

Fourteen pairs of occupational therapists (n = 28) took part in this study and tested 14 participants with a primary diagnosis of stroke. One pair was not able to complete the assessment leaving 13 sets of completed data. The therapist **test administrators** had qualified between 1964 and 1991, and comprised of five basic grades, five senior II, two senior I, one head IV and one deputy head occupational therapist. Nearly half of the therapists were working in "geriatric" or "care of the elderly" settings (n = 6). The other therapists encountered elderly patients as part of their case load on neurology or medical and surgical wards. Therapists' experience with older patients ranged from less than 1 to 15 years: less than 1 (n = 2), 1 to 5 (n = 6), 6 to 10 (n = 1) and 11 to 15 (n = 3). The distribution for experience with stroke patients was similar: less than 1 (n = 2), 1 to 5 (n = 8), 6 to 10 (n = 1), 11 to 15 (n = 2). Eleven of the therapists had known the patient prior to the research. Pre-test intervention comprised of informal observation (n = 3), assessment (n = 4) or assessment and treatment (n = 4). Five therapists mentioned that they had previously administered an ADL assessment, two had carried out motor assessments, one had undertaken a sensory assessment, three patients had been cognitively assessed and three therapists had carried out perceptual assessments.

The therapist **observers** had qualified between 1967 and 1990, and comprised of three basic grades, six senior II, one senior I, two head IV, one head III and one occupational therapist of unspecified grade. Five of the therapists were working in geriatric or care of the elderly settings and the other therapists were based in medical, neurology, orthopaedics, rheumatology, outpatient and day hospital settings. Therapists' experience with older patients ranged from less than 1 to 15 years: less than 1 (n = 5), 1 to 5 (n = 7), 6 to 10 (n = 1) and 11 to 15 (n = 1). The distribution for experience with stroke patients ranged from less than one to 10 years: less than 1 (n = 3), 1 to 5 (n = 8), 6 to 10 (n = 2). Seven of the observing therapists had known the patient prior to the research. Intervention comprised of informal observation (n = 1), assessment (n = 1) or assessment and treatment (n = 4). Two therapists mentioned that they had previously administered an ADL assessment, one had carried out a motor assessment, one had undertaken a sensory assessment, one patient had been cognitively assessed and two therapists had carried out perceptual assessments.

Of the 14 participants who took part in this study, eight had Right Hemisphere Lesions resulting in left hemiplegia, four had Left Hemisphere Lesions resulting in right hemiplegia, and two had strokes of unspecified type. The time between onset of stroke and testing ranged up to three months: less than one month (n = 6), 1 to 2 months (n = 4), 2 to 3 months (n = 4). Secondary diagnoses varied with the most common being hypertension, diabetes or arthritis. Two participants had a history of previous stroke. The locations used for testing included: occupational therapy departments (1st test n = 6, retest n = 4); wards (1st test n = 3, retest n = 3); day hospitals (1st test n = 2, retest n = 2); rehabilitation units (1st test n = 1, retest n = 1); an activity unit (1st test n = 1, retest n = 1); a rehabilitation therapy area (1st test n = 1, retest n = 1); and a research room (1st test n = 0, retest n = 1).

Description of sample and testing situation for the second study

The first author and three occupational therapy research assistants (one basic grade, one senior II, and one head occupational therapist) collected the data for the second study. Twenty-three participants were tested and the sample comprised of participants with the following primary diagnoses: stroke (n = 7); dementia (n = 15) and head injury (n = 1).

Summary description of the combined sample

Data from the two studies was combined for the statistical data analysis. The overall sample was comprised of 32 occupational therapists (covering all grades from basic to head occupational therapist) and 37 participants (with primary diagnoses of: 21 stroke; 1 head injury; and 15 dementia). The participant sample contained 54.1 percent (n = 20) females and 45.9 percent (n = 17) males aged between 60 and 91 years (Mean 75.6, s.d. 8.2).

Description of Statistical Analysis

At the time the studies were conducted (1991-1992) there was debate in the field of occupational therapy concerning the 'correct' statistic to use to estimate test-retest and inter-rater reliability. Ottenbacher and Tomchek (1993), reviewed 20 articles (from the American Journal of Occupational Therapy and Physical Therapy), which reported reliability studies. Amongst the statistics discussed in their paper, those suitable for the type of data collected in this study were Kappa, chi-square, and percent agreement. Ottenbacher and Tomchek concluded that Kappa was one of "the preferred methods of computing reliability in applied environments" (p. 14); Kappa was preferred to percent agreement as it corrects for chance agreement. Discrepancies were found between the average Kappa values and the average percentage agreement indexes evaluated in their study; all the reliability coefficients in their study had a ceiling value of 1.00 or 100 percent, Kappa had an approximate average value of 0.5 compared to Percent agreement which had an approximate average of 0.75 (75 percent). It was, therefore, decided to compute several statistics for this study in order to compare the values obtained and examine whether the same items exhibit substantial differences in levels of reliability when reliability coefficients are calculated by the different statistical methods. All analyses were calculated using SPSS/PC+ software (Norusis, 1991). The statistical analyses undertaken for this study were: (1) Percentage agreement; (2) Pearson's chi-square, Fisher's exact test and Phi Coefficient; and (3) Cohen's Kappa. For all the analyses data, from the two test administrations or for the two raters, for each variable, was cross-tabulated in a two by two contingency table.

Results

Percentage agreement (P): Percentage agreement (P) is an expression of the probability of a consistent decision (Crocker and Algina, 1986). P is the simplest measure of consistency for mastery decisions and can be defined as the proportion of people consistently classified as either master-master (able-able) or nonmaster-nonmaster (unable-unable) using two criterion referenced measurements. A new variable was constructed by assigning any subject who was consistently classified a

value of one and inconsistently classified data a value of zero. P equaled the sum of these values divided by the maximum possible value of this sum (which can only be obtained if all decisions are consistent). P was then expressed as a percentage (Crocker and Algina, 1986). Some of the data in this study lacked variance; this resulted in the formation of one-by-one or two-by-one contingency tables. Addition statistics could not be calculated for these tables. As a result, percentage agreement was the only statistic that could be calculated for all test items, and was the value used to provide an estimate of the overall reliability of the SOTOF, the reliability of each of the items in the five sub-tests: i.e., Screening Assessment, Eating Task (Task 1), Washing Task (Task 2), Drinking Task (Task 3), and Dressing Task (Task 4) and the reliability of each of the items on the Neuropsychological Checklist. Detailed results of the analysis for each item can be found in Laver's (1994) PhD thesis (Appendix 14 Tables 14.1 to 14.5). Two summary tables below (Tables 4.1 and 4.2), show the range of values and average value for each of the five sub-tests. The average percent agreement for test-retest reliability for the SOTOF was 91.8 percent (range 89.5-97.7 percent). The average percent agreement for inter-rater reliability was 93.1 percent (range 90.3-97.5 percent). The highest average values for both types of reliability were obtained for the Screening Assessment.

Table 4.1: The Average percent agreement for test-retest reliability for the SOTOF

Sub-test	**Range of % agreement across all sub-test items**	**Average % agreement for sub-test**
Screening Assessment	96.3% - 100%	97.7%
Task 1	33.3% - 100%	90.3%
Task 2	50.0% - 100%	89.5%
Task 3	72.4% - 100%	90.1%
Task 4	77.8% - 100%	91.6%
	Average % agreement for SOTOF	**91.8%**

Table 4.2: The Average percent agreement for inter-rater reliability for the SOTOF

Sub-test	**Range of % agreement across all sub-test items**	**Average % agreement for sub-test**
Screening Assessment	90.0% - 100%	97.5%
Task 1	28.6% - 100%	93.8%
Task 2	60.0% - 100%	92.8%
Task 3	63.6% - 100%	90.9%
Task 4	57.1% - 100%	90.3%
	Average % agreement for SOTOF	**93.1%**

An additional variable was constructed for the analysis of the reliability of the Neuropsychological Checklist. As the SOTOF is based on a progressive diagnostic clinical reasoning process, it was considered possible that therapists might reach the

same decisions but from the observation of different tasks. It was, therefore, important to consider not just whether a specific deficit was recorded on the Neuropsychological Checklist under a specific sub-test heading, but whether raters identified the same deficits from the complete administration of the SOTOF. The new variable was constructed by giving a value of 1 (deficit present), to a participant whenever a deficit had been recorded in the Neuropsychological Checklist under the heading of at least one of the sub-tests and a value of 2 (deficit absent) when the deficit had not been recorded under any of the sub-test headings. Percentage agreement values for the Neuropsychological Checklist for each item can be found in Laver's (1994) PhD thesis (Appendix 14, Tables 14.6 to 14.11) and are summarised below in Tables 4.3 and 4.4. These tables show the range of values and average value for each of the five sub-test headings on the checklist (Screen, Tasks 1, 2, 3, and 4). The average percent agreement for test-retest reliability for the SOTOF Neuropsychological Checklist was 95.2 percent (range 92.4-97.6 percent). The average percent agreement for inter-rater reliability was 93.9 percent (range 90.5-96.6 percent). The combined test-retest percentage agreement for the SOTOF (sub-tests and Neuropsychological Checklist) was 93.5 percent. The combined inter-rater value was 93.5 percent as well.

Table 4.3: Average percent agreement for test-retest reliability for the SOTOF Neuropsychological Checklist

Sub-test	**Range of % agreement across all sub-test items**	**Average % agreement for sub-test**
Screening Assessment	82.4% - 100%	97.6%
Task 1	79.4% - 100%	97.2%
Task 2	88.2% - 100%	94.2%
Task 3	73.5% - 100%	94.6%
Task 4	76.5% - 100%	95.3%
Total	67.6% - 100%	92.4%
	Average % agreement for SOTOF	**95.2%**

Table 4.4: Average percent agreement for inter-rater reliability for the SOTOF Neuropsychological Checklist

Sub-test	**Range of % agreement across all sub-test items**	**Average % agreement for sub-test**
Screening Assessment	87.5% - 100%	96.6%
Task 1	79.2% - 100%	94.2%
Task 2	79.2% - 100%	93.5%
Task 3	79.2% - 100%	93.6%
Task 4	83.3% - 100%	94.6%
Total	75% - 100%	90.5%
	Average % agreement for SOTOF	**93.9%**

Chi-square, Fisher's exact test and Phi Coefficient: The null hypothesis for this analysis was that there was no relationship between the scores of the two raters or the scores from the two test administrations. **Pearson's chi-square** statistic was used to compare the observed score distributions to those that would be expected if the two variables (the two sets of test scores from inter-rater and test-retest studies), were independent. The reliable use of chi-square is dependent on sample size (Norusis, 1991; Spitznagel, 1991). Assumptions related to sample size with contingency tables are based on the expected frequencies (Portney and Watkins, 1993), whereby, "if some of the expected frequencies in a table are less than 5, the observed significance level based on the chi-square distribution may not be correct" (Norusis, 1991, p. 270). One way to counteract this problem is to collapse variables (Sigel and Castellan, 1988; Portney and Watkins, 1993); however, as the contingency tables were already based on dichotomous variables it was not possible to combine variables to increase the expected frequencies in the contingency table cells.

Fisher's exact test: can be used to adjust chi-square to account for small expected frequencies and was calculated for this analysis. This test was used because it "evaluates the same hypothesis as the chi-square test, and it's suitable for tables having two rows and two columns for small expected frequencies" (Norusis, 1991, p. 270-271). Chi-square indicates if an association between variables is significant, the **Phi Coefficient** is used to express the degree of association between two nominal variables in a two-by-two table. The value of the Phi Coefficient ranges from -1.00 to +1.00 and can be interpreted as a correlation coefficient (Portney and Watkins, 1993). A significance level of 5% (< 0.05) was used to evaluate the significance of chi-square, Fisher's and Phi values. Values for these statistical computations were only available for a proportion of the sub-test and checklist items owing to a lack of variance. Summaries of results are shown below in tables 4.5 and 4.6. The full results can be found in greater detail in Laver's (1994) PhD thesis (Appendix 14, Tables 14.1 to 14.11). Table 4.5 shows the total number of items that were significant at the <0.05 level (Pearson's Chi-square, Phi and Fisher's exact test) for test-retest and inter-rater reliability for the Screening Assessment, Eating Task (Task 1), Washing Task (Task 2), Drinking Task (Task 3), and Dressing Task (Task 4). Those items that were not significant at this level fall into three categories. First, it was not possible to calculate these statistics for all test items as some of the two-by-two contingency tables contained missing data values. Second, some items were significant at the <0.05 level for Pearson's Chi-square and Phi but not for Fisher's exact test (two sided probability). Third some items were not significant at the <0.05 for any of the statistical tests. A breakdown of the analysis for each test item is in Laver's (1994) PhD thesis (Appendix 14, Tables 14.1 to 14.5).

Table 4.5: Significance of inter-rater and test-reliability for the SOTOF Screening Assessment and Four ADL Tasks

SOTOF component	Test-retest reliability: Number of significant items expressed as a fraction of the total number of items in that Task	Inter-rater reliability: Number of significant items expressed as a fraction of the total number of items in that Task
Screening Assessment	8/9	8/9
Eating Task (Task 1)	9/26	10/26
Washing Task (Task 2)	11/27	9/27
Drinking Task (Task 3)	6/28	10/28
Dressing Task (Task 4)	12/19	11/19

Results varied from item to item. All the items on the Screening Assessment were significantly related at the <0.05 level for both inter-rater and test-retest reliability, except for one item each that did not produce a two-by-two table. Only seven of the 26 items on Eating Task (Task 1: Eating from a Bowl using a Spoon) were not significantly related at the <0.05 level for test-retest reliability, and only three of the items on the Eating Task were not significantly related for inter-rater reliability. A similar distribution emerged for Washing Task (Task 2: Washing Hands in a Bowl): seven of the 27 items were not significantly related at the <0.05 level for test-retest reliability, and only two items were not significantly related for inter-rater reliability. For the Drinking Task (Task 3: Pouring and Drinking) seven of the 28 items for test-rest and three items for inter-rater reliability were not significantly related. In Dressing Task (Task 4: Putting on a Shirt), only two of the 19 items for test-retest and only one item for inter-rater reliability were not significantly related at the <0.05 level. Overall, the results indicated that the majority of items showed agreement across raters and, to a lesser extent, across time.

A pattern emerged for some types of items, from the four tasks, that were not significantly related at the <0.05 level. At least one of the "right / left discrimination" items was not significantly related for test-retest reliability on the first three tasks (Eating Task, Washing Task and Drinking Task). Patients rarely switch concepts of right and left completely but tend to exhibit general confusion in differentiating left from right. These items could have produced non-significant values because a deficit in right/left discrimination does not always result in a consistent response, but is more likely to appear as random performance with the subject sometimes placing the item correctly and sometimes giving an incorrect response.

The "recognition of objects" item was not significantly related for test-retest reliability in the Eating Task (Task 1), Washing task (Task 2) and Drinking Task (Task 3). The "describes use of objects" was also non-significant for test-retest reliability in three of the tasks (Eating Task, Drinking Task and Dressing Task). A possible explanation for these results could have been a learning effect if the subjects had been informed of the name and purpose of the objects by any of the raters during the first test administration. In clinical practice, therapists use assessment results as a starting point from which to educate patients. Raters in the second study had been trained by the researcher and did not offer such feedback. It was not possible to retrospectively

examine whether raters from the first study had given feedback to patients following the first test administration. Further research would be required to clarify this point.

The 'when handed' objects items were not significantly related at the <0.05 level for inter-rater reliability for all four tasks. This could have resulted from some ambiguity regarding both the administration and scoring of these items. This ambiguity came to light during the norming and was clarified in the original SOTOF test manual (Laver and Powell, 1995).

Other items that were not significantly related at the <0.05 level, appeared to be randomly distributed across tasks or only occurred in one of the four tasks. The test-retest reliability of the colour recognition items, for example, was significantly related for all but the Dressing Task (Task 4). The colours on the other three tasks could have been easier to perceive owing to the size of the objects and because brighter primary colours were used (the button used for the second study was dark blue). This problem might be solved by increasing the size of the button used and changing to an easily perceived colour, such as yellow or red. This would also address the problem of using dark colours from the blue/green end of the spectrum which are more difficult for older people to perceive owing to primary ageing which causes yellowing of the retina. There could have been a learning effect on this item if any of the raters had corrected the patient and informed them of the colour of the button during the first test administration.

The Screening Assessment is used to evaluate whether the person is functioning at the baseline level defined in the criteria for the administration of the SOTOF. Patients, therefore, should have passed the majority of the Screening Tasks if they had been entered in the rest of the study. Because of this high pass rate many of the deficits under the Screen heading of the Neuropsychological checklist lacked variance and statistics could not be computed for a large proportion of these items. (Percentage agreement for these items was very high ranging from 82.4% to 100% with an average of 97.6%). All of these items were significantly related for inter-rater reliability indicating considerable agreement among test administrators. All but two items were significantly related for test-retest reliability, these were expressive language and hearing acuity. Both these functions would not have been expected to alter in stroke patients during such a short space of time. The non-significant value obtained for the hearing acuity item is more likely to be the result of random errors; possible explanations include changes in the level of background noise in the testing environments or the failure of the participant to use a hearing aid (if required), during one of the two test administrations.

Summaries of the results for items on the Neuropsychological Checklist are provided in Table 4.6. The full detailed results can be found in Laver's (1994) PhD thesis (in Appendix 14, Tables 14.6 to 14.11). Table 4.6 shows the total number of items that were significant at the <0.05 level (Pearson's Chi-square, Phi and Fisher's exact test) for test-retest and inter-rater reliability for each Neuropsychological deficit under the five Checklist headings (Screen, Task 1, Task 2, Task 3, and Task 4) and the constructed "Total" variable. All values are presented as a fraction of the total number of Neuropsychological Checklist items for each deficit (i.e. out of a total of six items per deficit). Those items that were not significant at this level fall into the same three categories described above.

Table 4.6: Significance of Reliability of Neuropsychological Checklist Items

Deficit	Test retest: Number of significant items (maximum = 6)	Inter rater: Number of significant items (maximum = 6)
Language: comprehension	5	2
Language : expression	3	4
Hearing : acuity	3	1
Hearing : auditory agnosia	0	0
Cognition : orientation	1	1
Cognition : attention	5	0
Cognition : short term memory	1	2
Cognition : long term memory	1	0
Motor : abnormal tone	6	6
Sensation : proprioception	6	6
Sensation : tactile discrimination	2	0
Vision : acuity	0	0
Vision : Visual attention	0	0
Vision : visual scanning	0	0
Vision : visual field loss	0	0
Vision : visual neglect	2	0
Agnosia : visual spatial	0	0
Agnosia : visual object	1	1
Agnosia : colour agnosia	0	0
Agnosia : tactile agnosia	5	3
Apraxia : constructional	2	0
Apraxia : dressing apraxia	2	2
Apraxia : Motor apraxia	2	0
Apraxia : ideomotor apraxia	2	0
Apraxia : ideational apraxia	2	0
Body Scheme : somatognosia	0	0
Body Scheme : unilateral neglect	3	4
Body Scheme : anosognosia	0	0
Body Scheme : right / left discrimination	1	3
Spatial Relations : figure ground	3	0
Spatial Relations : position in space	3	1
Spatial Relations : form constancy	0	0
Spatial Relations : spatial relations	3	1
Spatial Relations : depth perception	0	0
Spatial Relations : distance perception	0	0
Perseveration :	0	0

Many of the Neuropsychological Checklist items did not produce statistical values owing to lack of variance; it should be noted that the majority of these items had a

percentage agreement of 100%. Some items were significantly related at the level <0.05 level but only for Pearson's Chi-square and Phi, not for Fisher's exact test.

A pattern emerged for some of the non-significant items, for example, the 'Sensation: tactile discrimination' was inconsistently recorded by raters across all four Task headings and as examined through the constructed 'Total' variable. Examination of the Task observational checklist assessment forms showed an inconsistency between raters regarding the scoring for the 'identifies object through touch-left hand' item, especially when the subject had previously identified the object with his/her right hand. The 'Agnosia: tactile agnosia' item, which is also identified through the performance of these 'identification through touch' items, were inconsistently recorded for both test-retest and inter-rater reliability.

The 'Cognition: short term memory' item was not significantly related for both test-retest and inter-rater reliability under the Eating Task (Task 1), Washing Task (Task 2) and Drinking Task (Task 3), Neuropsychological checklist headings. Both types of reliability, however, were significantly related when examined through the constructed 'Total' variable for this deficit. This suggests that the short term memory deficit is identified consistently overall by the test administrators, but does not manifest during any one specific Task performance. A similar pattern also emerged for 'Cognition: attention' which had significant values for both types of reliability for the Total variable despite non-significant inter-rater reliability values for the Eating Task (Task 1) and Drinking Task (Task 3), and a non-significant test-retest value for Washing Task (Task 2).

Other deficits that were not significantly related at the <0.05 level for some of the tasks but which were consistently recorded over the whole checklist as indicated by significant 'Total' values were: 'Language: expression'; 'Agnosia: visual object agnosia'; 'Apraxia: ideomotor apraxia'; 'Apraxia: ideational apraxia'; 'Body scheme: right/left discrimination'; and 'Spatial relations: spatial relations'. Conversely only three deficits produced non-significant values for the Total variable: 'Vision: visual attention' was non-significant for test-retest reliability, 'Agnosia: visual spatial' and 'Spatial relations: figure ground discrimination' were non-significant for inter-rater reliability. Deficits which had some items that were not significantly related for some Task headings and for the Total variable were: 'Language: comprehension'; 'Hearing: acuity'; 'Cognition: long term memory'; 'Vision: visual scanning'; 'Vision: visual field loss'; 'Apraxia: constructional apraxia'; and 'Perseveration'. The number of items that were not significantly related at the <0.05 level for these deficits ranged from two to six.

Cohen's Kappa (K): Cohen's Kappa (K) is a measure of agreement which has "been proposed for categorical variables [and] can be applied to an arbitrary number of raters" (Siegel and Castellan, 1988, p. 284). Kappa provides a transformation of P to a new scale in which the points 0 and 1 are interpretable:

> "where Pc is the chance probability of a consistent decision... that is, the probability for the hypothetical situation in which the scores on the two forms are statistically independent. Statistical independence of test scores implies that decisions are statistically independent. The coefficient Pc is sometimes referred to as the *chance consistency*... chance consistency can be viewed as

> a baseline for judging the actual amount of consistency observed for the two forms [administrations of the test]. Thus K may be interpreted as the increase in decision consistency that tests provide over chance expressed as a proportion of the maximum possible increase over chance consistency"
>
> (Crocker and Algina, 1986, p. 200-201)

Coefficient K is 0 when there is no increase and 1.0 when there is maximal increase. A value of 0 does not mean that decisions are so inconsistent as to render the item worthless, but that the decisions are no more consistent than decisions based on statistically independent scores. This consistency could still be substantial (a minimum of 50% (0.5) for exchangeable test forms). A value of 1 indicates that decisions are as consistent as those based on perfectly statistically dependent scores (Crocker and Algina, 1986; Siegel and Castellan, 1988; Norusis, 1990). "The coefficient K can assume negative values...which corresponds to the situation in which there is an inverse relationship between the scores on the two forms" (Crocker and Algina, 1986, p. 201). Kappa treats all inconsistent classifications as equally serious. As the SOTOF does not use a continuous scoring system or scale, statistics which evaluate the magnitude of the discrepancy of a misclassification in judging reliability of decisions were not relevant.

SPSS/PC+ was used to compute Cohen's Kappa with asymptotic standard error (ASE1) and the t statistic value. "The test of the null hypothesis that kappa is 0 can be based on the t statistic. The t value is the ratio of the value of kappa to its asymptotic standard error when the null hypothesis is true. [N.B.] the asymptotic standard error on the [SPSS/PC+] output does not assume that the true value is 0" (Norusis, 1990, p. 136-137). Full results for Kappa, ASE1, and t values for each of the SOTOF sub-test items and the Neuropsychological Checklist items can be found in Laver's (1994) PhD thesis (Appendix 14, Tables 14.12 to 14.22). Kappa values were only available for a proportion of the test and checklist items owing to lack of variance. Only one item in the entire test (test-retest reliability Washing Task 'continues action unnecessarily') obtained a value of zero which indicated that decisions were no more consistent than decisions based on statistically independent scores. The scoring of this item was identified as ambiguous during the Norming Study and was clarified in the original SOTOF test manual (Laver and Powell, 1995). Nine sub-test items obtained a value of one for test-retest reliability indicating that decisions were as consistent as those based on perfect statistically dependent scores. Fourteen sub-test items also obtained a value of one for inter-rater reliability. On the Neuropsychological Checklist 15 items had a value of one for test-retest reliability and nine for inter-rater reliability. It was impossible to obtain Kappa values for all test items, average Kappa values could only be calculated from a proportion of the items and should, therefore, be viewed as approximate values. Average Kappa values are shown in Tables 4.7 to 4.8.

Table 4.7: approximate average Kappa values for the Screening Test and four ADL Tasks for test-retest reliability

Sub-test	Number of SOTOF items that kappa could be calculated for	Range of Kappa values across sub-test items	Average Kappa value for sub-test
Screening Assessment	10 / 11	0.78 - 1	0.92
Task 1	17 / 26	-0.04 - 0.9	0.47
Task 2	19 / 27	-0.07 - 0.77	0.38
Task 3	12 / 28	-0.09 - 0.66	0.37
Task 4	15 / 19	-0.07 - 1	0.67
Average Kappa value for SOTOF	73 / 111	-0.09 - 1	0.56

Table 4.8: approximate average Kappa values for the Screening Test and four ADL Tasks for inter-rater reliability

Sub-test	Number of SOTOF items that kappa could be calculated for	Range of Kappa values across sub-test items	Average Kappa value for sub-test
Screening Assessment	8 / 11	0.65 - 1	0.94
Task 1	10 / 26	-0.4 - 1	0.77
Task 2	7 / 27	0.23 - 1	0.5
Task 3	8 / 28	0.25 - 1	0.61
Task 4	12 / 19	0.4 - 1	0.75
Average Kappa value for SOTOF	73 / 111	-0.4 - 1	0.71

The approximate average Kappa values for the Screening Test and four ADL Tasks ranged from 0.37 to 0.92 (average 0.56; see Table 4.7) for test-retest reliability and from 0.5 to 0.94 (average 0.71; see Table 4.8) for inter-rater reliability. The overall average Kappa value for test-retest reliability for the SOTOF was calculated from values available for 53.2% of items and was 0.56. The overall average Kappa value for inter-rater reliability for the SOTOF was calculated from values available for 40.7% of items and was 0.63. These values are slightly above the average Kappa value (0.5), reported by Ottenbacher and Tomcheck (1993) in their evaluation of reliability analysis in therapeutic research.

Table 4.9: Approximate average Kappa values for the Neuropsychological checklist for test-retest reliability

Checklist Sub-test Heading	Number of SOTOF items that kappa could be calculated for	Range of Kappa values across sub-test items	Average Kappa value for sub-test
Screening Assessment	7 / 36	0.21 - 1	0.63
Task 1	18 / 36	-0.04 - 1	0.56
Task 2	18 / 36	-0.05 - 0.67	0.44
Task 3	15 / 36	-0.06 - 0.84	0.47
Task 4	16 / 36	-0.05 - 1	0.61
Total	27 / 36	-0.04 - 1	0.59
Average Kappa value for SOTOF	/ 216	-0.06 - 1	0.55

Table 4.10: Approximate average Kappa values for the Neuropsychological checklist for inter-rater reliability

Checklist Sub-test Heading	Number of SOTOF items that kappa could be calculated for	Range of Kappa values across sub-test items	Average Kappa value for sub-test
Screening Assessment	6 / 36	0.47 - 1	0.8
Task 1	14 / 36	-0.09 - 1	0.52
Task 2	13 / 36	-0.06 - 1	0.52
Task 3	13 / 36	-0.11 - 1	0.47
Task 4	17 / 36	-0.04 - 1	0.5
Total	25 / 36	-0.07 - 1	0.44
Average Kappa value for SOTOF	88 / 216	-0.09 - 1	0.54

Comparison of the values obtained by each of the statistical analyses

Summary tables for the three analyses were constructed for the items on the Neuropsychological Checklist to examine the discrepancy of reliability values obtained through each of the statistical methods and can be found in Laver's (1994) PhD thesis (Appendix 14, Tables 14.23 to 14.26). Comparison of percentage agreement and Kappa values obtained in this study supported the finding by Ottenbacher and Tomchek (1993) that percentage agreement values were

consistently higher than kappa values. Average percentage agreement for test-retest reliability of the SOTOF (calculated from values for all items) was 0.94 (93.5%), compared with an approximate (calculated from values available for only 53.2 percent of items) average Kappa value of 0.56. Average percentage agreement for inter-rater reliability of the SOTOF (calculated from values for all items), was also 0.94 (93.5%) compared with an approximate (calculated from values available for only 40.7 percent of items), average Kappa value of 0.63. Comparison of Kappa values with the significance level of values obtained by Chi-square, Fisher's and Phi showed that items with Kappa values of 0.5 and above were usually significant (at the <0.05 level) for these other analyses. Items with Kappa values between 0.34 and 0.65 were significant for Chi-square and Phi but did not always produce significant values for Fisher's exact test. Items with Kappa values less than 0.34 usually had non-significant values for the three other statistical analyses.

Conclusions

The use of Cohen's Kappa, Chi-square (adjusted for small sample sizes where necessary) and Phi Coefficient produce more conservative estimates of reliability than Percentage agreement and are, therefore, preferred methods of analysis. The Kappa value is easy to interpret and gives the advantage of accounting for chance agreement; the results of this study supported Ottenbacher and Tomchek's (1993), recommendation of Kappa as a preferred method of computing reliability in applied therapeutic research. Unfortunately, a lack of variance in some of the data meant that Kappa could not be calculated for all the SOTOF test items. The average Kappa values are, therefore, only approximations of the overall reliability. It was necessary to rely on Percentage Agreement values, however, they should be treated with some caution as they may give an over positive image of the test's reliability.

The Screening Assessment appears to have very good test-retest (97.7 percent, Kappa approximate value of 0.92), and inter-rater reliability (97.5 percent, Kappa approximate value 0.94), and can be used as a reliable indication of gross motor, visual and cognitive functioning. The four ADL Tasks have higher inter-rater reliability (90.3-93.8 percent, Kappa: 0.5-0.77) than test-rest reliability (89.5-91.6 percent, Kappa: 0.37-0.67). This could be the result of genuine fluctuations in subjects' performance over the two administrations of the test. The research assistants who conducted the testing for the second study noted what they considered to be genuine changes in the performance of some participants with dementia from one test administration to another. A few of the occupational therapists who conducted the first study noted changes in the performance of some of their stroke patient subjects. This was partly the result of participants responding to therapists' corrections during the first test administration (e.g. learning a hemiplegic dressing method shown during the first test enabled independent dressing in the second test), and to perceived changes in function from one day to the other. Patients in the early stages following stroke can make spontaneous recovery. Furthermore the rationale behind practice of ADL tasks in occupational therapy is based on the belief that repetition of tasks aids the return of function. The fact that the re-test was a repetition of task performance could have also resulted in some slight increase in functional performance.

Both the average percent agreement and Kappa values for the SOTOF are higher than the average of the values reported in the reliability studies evaluated by Ottenbacher and Tomchek (1993): SOTOF's test-retest average of 91.8 percent and inter-rater average of 93.1 percent were higher than the average values for these 20 studies which was approximately 75 percent; average Kappa values of 0.56 for test-retest and 0.71 for inter-rater reliability were also higher than the 0.5. average value reported for these studies.

As the Neuropsychological Checklist score is based on diagnostic reasoning and requires rater judgement, it was anticipated that its reliability would be less than the SOTOF Tasks, and lower than other Neuropsychological Assessments. However, the average percent agreement for test-retest reliability was 95.2 percent and the approximate average Kappa value was 0.55. Inter-rater reliability was very similar at 95.2 percent / 0.54. These figures are encouraging, particularly when the ranges of experience of the clinicians used in this study are taken into consideration.

The average percentage agreement and approximate average Kappa values obtained on the SOTOF's sub-tests and Neuropsychological Checklist compared favourably to other occupational therapy standardised assessments available at the time of the study (early 1990s). The SOTOF values were particularly encouraging in light of the fact that the test involves a major component of rater judgement (therapist’s clinical reasoning). This supported the supposition that observation of a patient's performance in ADL tasks can provide as reliable a picture of neuropsychological deficit as the more formal psychological test batteries currently in use.

5 Examination of the validity and clinical utility of SOTOF (1st edition)

Summary

Two studies were undertaken to examine aspects of the content validity, face validity and clinical utility of SOTOF. Three sample groups (qualified occupational therapists, occupational therapy students and patients with a primary diagnosis of stroke) contributed to the studies. Data was collected via two self-administered questionnaires and one administered questionnaire. Forty-four occupational therapists, forty patients and thirty-three students completed the surveys. Review of the content of SOTOF, by qualified occupational therapists, identified that SOTOF addresses the eight constructs of perceptual function, sensory function, motor function, cognitive function, language, performance of Activities of Daily Living (ADL), visual function and auditory function. Overall SOTOF appears to have very good clinical utility as it was easily understood, easily administered and quick to use. Therapists found the materials were easy to obtain, carry, clean and store. The test was perceived, by both qualified occupational therapists and students, to be relevant for the patient group for whom it was designed. Therapists reported it was suitable for use by all qualified occupational therapists and occupational therapy students under supervision. They also reported that the test was not stressful for patients. SOTOF appeared to have good face validity with the patients for whom the test was designed. The majority of subjects felt SOTOF tested their ability and/or function. The test appears to be relevant for the population. None of the subjects reported minding being asked to do the SOTOF tasks. Ninety-five percent thought the tasks were things they would normally do, only 15% found the test irrelevant. SOTOF does not appear to induce test anxiety in the majority of test takers; only 12.5% reported finding the test stressful. This chapter concludes with an examination of training issues and potential alterations to the SOTOF.

Introduction

Validation is undertaken by a test developer to collect evidence to support the types of inferences that are to be drawn from the results of a test (Crocker and Algina, 1986). There is no one recognised measure of validity. It is usual to conduct a range of studies to examine its different aspects (Bartram, 1990). Three types of validation, content, construct and criterion-related validation, are traditionally performed (Crocker and Algina, 1986). The first study to evaluate SOTOF addressed aspects of its content validity, face validity and clinical utility, as perceived by occupational therapist test administrators and stroke patients taking the test.

Defining the Concepts of Content and Face Validity

Content validity refers "to the degree to which a test measures what it is supposed to measure judged on the appropriateness of the content" (Bartram, 1990, p. 77). It

depends "on the relevance of the individual's test responses to the behaviour area under consideration, rather than on the apparent relevance of item content" (Anastasi, 1988, p. 140). Therefore, a content validation study should involve the actual administration of the test, as opposed to analysis of content solely from the test manual and materials. Content validity is usually judged by the professional group who are to use the test and/or by a panel of experts (Bartram, 1990; Crocker and Algina, 1986). "The purpose of a content validation study is to assess whether the items adequately represent a performance domain or construct of specific interest…" (Crocker and Algina, 1986, p.218). It involves the systematic analysis of the behavioural domain to check whether "all major aspects of the domain are covered by the test items, and in the correct proportions" (Anastasi, 1988, p. 140). This requires the domain of concern to be clearly defined in advance (Anastasi, 1988; Crocker and Algina, 1986). This study focused on the evaluation of SOTOF by a group of therapists, representative of those test administrators for whom the test was designed. The study included an evaluation of both the constructs and behaviours addressed by the test.

Face validity "is the dimension of a test by which it appears to test what it purports to test" (Christiansen and Baum, 1993, p 851). The concepts of content and face validity are similar, but should not be confused (Anastasi, 1988). "The difference is that face validity concerns the acceptability of a test to the test-taker, while content validity concerns the appropriateness of the content of the test as judged by 'professionals'…" (Bartram, 1990, p.77). All definitions of face validity agree that the test should be acceptable to the test-taker. However, definitions of face validity vary in terms of who else the test should appear to be acceptable. Bartram (1990) defined face validity as solely "the degree to which the test-taker sees a test as being reasonable and appropriate" (p. 76). Crocker and Algina (1986) broaden this definition to include "laypersons or typical examinees" (p. 223). Anastasi (1988) perceives face validity as pertaining "to whether the test 'looks valid' to the examinees who take it, the administrative personnel who decide on its use, and other technically untrained observers" (p. 144). For the purpose of this study, face validity was examined simply in terms of the perceptions of the subjects who took the test.

As face validity is not validity in the technical sense it has little direct psychometric importance. However, its evaluation was considered to be important for several reasons. Anastasi (1988) draws attention to literature in the field of psychometrics which highlight "the paucity of available research on face validity, despite its probable contribution to prevalent attitudes towards tests" (p. 145). Within the field of occupational therapy there has been recent criticism regarding the use of tests based on items which have little meaning and relevance for the test taker (Fisher and Short-DeGraff, 1993; Trombly, 1993; Law, 1993). Yet, none of the test manuals, or articles on developed occupational therapy tests, reviewed in the early 1990s during the development of the SOTOF addressed face validity in any way. Respondents of a questionnaire, to evaluate occupational therapists' assessment practice and identify the criteria and format for SOTOF, emphasised the need for tests to be perceived as appropriate and meaningful for their patients. Good face validity can have indirect effects on the outcome of a subject's performance "by facilitating rapport between the test and the test-taker which may, in turn, increase reliability" (Bartram, 1990, p. 76). Volition is important because "people are more

likely to take seriously activities which seem reasonable and which they feel they understand" (Bartram, 1990, p. 76).

Defining the Concept of Clinical Utility

It is important to select the most suitable test for a particular patient (Christiansen, 1993; Opacich, 1992; Bonder, 1990; Kline, 1990; Law and Letts, 1989). Test critique involves not only an examination of the standardisation and psychometric properties of a test but also its clinical usefulness, or utility. The terms clinical usefulness and clinical utility appear to be used interchangeably in the majority of cases. However, Murdock (1992a) perceived clinical utility to be an aspect of clinical usefulness. She viewed the concept of clinical usefulness as comprising four elements, format, cost, acceptability and utility. Feinstein, Josephy and Wells (1986) used the term "clinical sensibility" (p. 413) to address similar issues. Drawing upon the work of several authors (Christiansen, 1993; Murdock, 1992a; Law and Letts, 1989; and Feinstein, Josephy and Wells, 1986), it was decided that the concept of clinical utility would be applied in the broadest of sense for this study and encompass “the issues of instructions, cost, time, acceptability, and format" (Law and Letts, 1989, p. 524).

In a global sense, cost encompasses issues of financial outlay, time and energy. "Costs are reflected in the amount of equipment, space, training time, and expertise necessary to administer and interpret an assessment" (Christiansen, 1993, p. 258). Energy relates to "the ease with which an index can be used and analyzed" (Feinstein, Josephy and Wells, 1986, p. 418). The format of the test and the clarity of instructions will have a direct bearing on the ease in which a test is understood and administered. The test must be acceptable to the therapist and should "also be acceptable to the client and to his or her family, who should understand and agree with the usefulness of the items being measured" (Law and Letts, 1989, p. 524). Acceptability is encompassed in the concept of face.

To include all relevant factors in the evaluation of the clinical utility of SOTOF the areas addressed by this study were not only identified by literature review, but also by revisiting the criteria identified by the respondents of the questionnaire (to evaluate occupational therapists’ assessment practice and identify the criteria and format for SOTOF). The content validity, face validity and clinical utility of SOTOF were examined by samples of occupational therapists and stroke patients through a postal survey in 1991.

Methodology

The 1991 study involved two related sample groups; volunteer hospital based occupational in the United Kingdom and patients (participants) who had a primary diagnosis of stroke, and were recruited from these therapists' caseloads. The therapists studied the SOTOF manual, administered the test to at least one person with stroke, and then completed a survey which asked questions about the content, utility and relevance of the test. The patients undertook the SOTOF and answered questions, concerning their opinion and experience of the test, during a structured interview.

Research Questions

Questions pertaining to **Content Validity** explored therapists' perceptions of the constructs and behaviours addressed by SOTOF and the neuropsychological deficits which SOTOF could be used to highlight in elderly stroke patients. Questions pertaining to **Clinical Utility** related to the test manual, test materials, length of test administration, appropriateness of the test to the patient group, test induced anxiety, and level of expertise required to administer the test. Questions pertaining to **Face Validity** addressed patients perceptions of the purpose of SOTOF, what was tested, whether the SOTOF tasks were activities the patients would normally engage in and whether patients minded being asked to do the SOTOF tasks? The patients' experiences of taking SOTOF (for example levels of experienced stress, boredom, and enjoyment) were explored.

Procedure for sampling the population

Two samples were recruited for this study in pairs, an occupational therapist (test user) and a patient (recipient) group. To increase homogeneity it was decide to limit the diagnostic criteria for patients to a single diagnosis. Stroke was selected as it is a common condition, is relatively well recognised and is known to cause neuropsychological deficits. The second criterion for participant recruitment was their age. SOTOF was developed for an older client group. Subjects had to be 60 years or over to be recruited for the study. Although, the most usual age criteria for access to gerontology services within the NHS is sixty-five, a cut-off of sixty years was selected for this study as some units do take patients who are a few years younger. One of the aims of SOTOF was to develop a test that could be administered by all qualified occupational therapists regardless of their grade or experience, and without the need for additional training. The criteria for therapists was limited only by their access to the patient group: therapists had to be working with stroke patients aged sixty and above.

The provisional studies for the development and piloting of SOTOF had been undertaken in a small urban area. It was decided to broaden the geographical limits of this study to obtain a more representative sample. The study was conducted at a national level across the whole of the United Kingdom (U.K.) and Ireland. It would have been too expensive to approach randomly selected subjects directly by post or telephone, thus a volunteer group of therapists was used. Volunteers were preferred because of the time consuming nature of the study. It was felt that the amount of work the subjects were expected to do (read SOTOF test manual, select a patient, administer the test, interview the patient and complete a questionnaire) would result in a high drop-out and non-response rate. In addition, the cost of test and survey materials was very high. It was too expensive to replace drop-outs unless materials were returned unused; therefore, obtaining motivated and interested participants at the outset of the study was of paramount importance.

To attract volunteers it was decided to place a letter in The British Journal of Occupational Therapy (BJOT) as it has a wide readership in the U.K. Even therapists who do not have individual membership often have access to the journal through their departments and colleagues. It was felt that therapists who read the journal were likely to be interested in new developments in the field and would be a

potentially motivated research group. Participants were recruited more informally when therapists expressed interest in the research at BAOT sub-group and regional meetings and at lectures and papers given by the researcher. A postal survey method was used for this study. Bennett and Ritchie (1975) stated that a 40-60 per cent rate of return is to be expected when conducting postal surveys. It was decided to survey at least one hundred therapists with a view to obtaining a sample size of about fifty subjects.

The development of the questionnaires

Therapists participating in the study were to administer the SOTOF to people with stroke. A method needed to be identified for collecting data from both therapist and patient participants. A wide geographically based sample was desired. The expense of travelling to interview subjects would have been prohibitive so methods of data collection were restricted to either a postal survey or telephone interviews. A postal survey was selected for several reasons: a telephone interview would not have been a suitable data collection method for patients; the amount of data to be collected from therapists would have required lengthy, and costly, telephone calls; and it was felt that a written, rather than a verbal, response format would provide the therapists with more opportunity for reflection. Two questionnaires were developed, one for the occupational therapist sample and one for the patient sample. The therapist questionnaire was to be self-administered. However, individuals frequently experience language, motor and visual deficits following stroke and it was felt that a considerable proportion of the patient sample would be unable to complete a self-administered questionnaire. The patient survey, therefore, took the format of an administered, structured interview with the interviewer recording the person's responses on the questionnaire.

The process of constructing medical questionnaires has been explained by Bennet and Ritchie (1975). Their text was used to guide the format and distribution of the questionnaires. A particular concern for a researcher developing questionnaires is the effect of bias:

> "Bias... is the intrusion of any unplanned or unwanted influence. It may occur firstly through inappropriate wording of questions, for example the use of leading questions and loaded words... the respondent himself is biased in the direction of his own self-interest... he may also cheat because of disinterest in the questionnaire or to please the interviewer."
>
> (Bennet and Ritchie, 1975, p. 23)

Two forms of bias are the response set and the halo-effect. A response set is "the tendency of a respondent persistently to respond in a certain way, irrespective of the question" (Bennet and Ritchie, 1975, p. 25). A positive halo-effect is a tendency to over-estimate qualities which the respondent perceives as desirable owing to a feeling of approval towards the interviewer or towards the thing to which the questions refer (Bennet and Ritchie, 1975). A major problem encountered by researchers using a postal survey method is "the very poor response rate they tend to produce; 40-60 per cent is a common rate of return" (Bennett and Ritchie, 1975, p. 56). Unreturned questionnaires not only affect the sample size but can also bias the sample to an unknown degree. Two methods for controlling response bias are correction and prevention. The preventative method involves constructing a

questionnaire in a manner which will yield the most returns. Factors such as the colour of paper used, format of questions, quality of print, questionnaire length, provision of stamped addressed envelopes, denomination of stamps used, whether the envelope should be handwritten or printed, the tone of the accompanying letter, layout and method of recording responses, amount of space provided for qualifying statements, and the time of year in which the questionnaire is sent should be considered. Research indicates that higher responses are gained "by using stamped, handwritten reply envelopes as opposed to printed machine-stamped envelopes which may be associated with advertising (Bennet and Ritchie, 1975, p. 57). Another method for reducing non-response is to follow up non-respondents with a telephone call or second copy of the questionnaire.

Three components of questionnaires were addressed in the two questionnaires developed for this survey:

1. the content, e.g., demographic, personality characteristics, behavioural patterns, health history;
2. the form of the question, e.g., forced-choice or open ended; and
3. the level of data collected, e.g., factual or attitudinal.

(Stein, 1989, p. 118)

The Occupational Therapist's Questionnaire

Self-administered questionnaires have been designed in a variety of forms, including computer administered and pen and paper formats. The later was chosen as a postal distribution method had been selected. The questionnaire took the format of a standardized printed form. The content of the therapist self-administered questionnaire (for copy see Laver's, 1994 PhD thesis, Appendix Four) comprised an introductory letter and a questionnaire containing twenty questions. One of the disadvantages of a postal questionnaire is the complete loss of contact with the respondent. An introductory letter can be used to compensate for this loss of rapport and is important for engaging the participant's interest and motivation. A postal questionnaire:

> "…should begin with an introduction for the respondent, written on a separate page to allow him to read it before glancing at the questions... it should give some explanation about the purpose of the questionnaire. This must be very general or it may introduce bias. The introduction should include the name of the relevant organizing body"
>
> (Bennet and Ritchie, 1975, p. 26)

Introductory information was provided in the letter on a separate sheet attached to the front of the questionnaire. The letter described the purpose of the survey, gave the researcher's name and address, and indicated a deadline for the return of questionnaires. The questionnaire consisted of four printed A4 sheets. Several types of questions were used. Both attitudinal and factual information was sought. Factual data included: details about the therapist (grade, number of years' experience, clinical area); details about the patient(s) tested for the survey (sex, age, diagnosis); and information about the length of time taken to administer SOTOF. Attitudinal data involved the therapist's perceptions of the content, administration, and relevance of the test. Both open-ended and forced-choice question formats were used. Questions

were designed to yield both qualitative and quantitative data. Rating scales were used for ten of the questions. These took a descriptive scale format and required the rater to choose a phrase from a list of five phrases. Two of the questions used a forced choice, dichotomous decision (yes/no) format. The therapist was invited to make additional comments at the end of the questionnaire.

The layout of a questionnaire is important as the visual impact can increase co-operation and help the respondent to work through the questions in a logical manner. "Questions should be clarified by the use of lined insets to help guide the interviewer or respondent through the form so that he will not write responses in the wrong place or omit subsidiary questions" (Bennet and Ritchie, 1975, p. 27). Lines were provided to guide the position of responses. Clear instructions were given to indicate response format (for example: please circle your choice, please list below, and please tick the boxes to indicate your choices). The use of coloured paper has been suggested for questionnaires as it can be perceived as attractive by respondents, and will stand out against other papers on the respondent's desk. Light colours should be used to show off the print to best advantage and yellow backgrounds have been found to have the highest percentage of returns in a postal survey (Bennett and Ritchie, 1975). The therapist questionnaire was printed onto light orange-yellow paper. This helped the questionnaire to stand out from the other information sent in the survey package.

The Patient Face Validity Questionnaire

The patient questionnaire was to be administered in a structured interview by a member of the multi-disciplinary team caring for the patient. It was felt that patients might find it easier to provide critical feedback to a professional other than the therapist who had administered the test to them. There was a disadvantage to this methodology, as the questionnaire was distributed and returned by post it was not feasible to train and supervise interviewers, nor monitor the patient-interviewer interaction. The questionnaire, therefore, had to be very clearly structured and provide clear questions to be used during the interview. There are two main types of administered questionnaires, standardised and unstandardised. There are advantages and disadvantages to both these methods. The unstandardised format allows the interviewer to reword items and introduce probes which can increase the subject's comprehension and response. However, such flexibility also "allows greater opportunity for the interviewer to introduce his own attitudes and opinions thus providing an important source of bias" (Bennet and Ritchie, 1975, p. 39). A standardised format was selected for this study:

> “The standardized interview has the advantage of constituting a standard instrument of measurement which allows for comparison of information derived from various sources and gives good results on test-retest repeatability… the standardized questionnaire is based on a series of questions which have been determined in advance of the interview, and which are asked in the same order and with the same wording for every respondent. The interviewer is allowed no initiative with regard to the introduction of additional items or probes which might be relevant to the individual case… his role is strictly limited by the content of the questionnaire”
>
> (Bennet and Ritchie, 1975, p. 38)

The layout of an administered questionnaire is just as important as the layout of a self-administered questionnaire "as bad design may confuse the interviewer, and result in the omission of items, administration of items out of order or misreading of instructions" (Bennet and Ritchie, 1975, p. 26). Both the layout and the content of the questionnaire were carefully selected. The questionnaire consisted of three printed A4 sheets. The questionnaire (for copy please see Laver's (1994) PhD thesis, Appendix Five) comprised an introductory section to be completed by the interviewer and seven questions to be answered by the patient. This introductory section involved the collection of factual data; it requested the patient's details (i.e. age, sex, diagnoses) and an indication of the parts of SOTOF which had been administered. After the section on patient details, a statement was given to the patient as an explanation of the purpose of the questionnaire. This was followed by a series of questions. The level of data collected through the seven questions was attitudinal and pertained to the patient's opinions about the content, nature and purpose of SOTOF and his / her experience of undertaking the test. The questions were numbered, printed in bold, and kept as concise as possible. Lined spaces were given after each question to provide the interviewer with a clear indication of where to record results. Six of the questions (numbers 1 - 5 and 7) were open ended and resulted in qualitative, descriptive data.

To discourage an acquiescent response set or a positive halo-effect, question six was designed in a different format. It was felt that older patients might find it easier to give positive rather than negative feedback about the test. Issues of relevance, stress and difficulty were listed for the patient. Five pairs of words (for example: easy/difficult; boring/interesting) which could describe an individual's experience of the test were selected. The interviewer attached each of these words to the question "Did you find the assessment... [e.g. easy]?" A forced-response, dichotomous categorical (yes/no) format was used for these questions. It was felt the provision of terms such as stressful, difficult and boring would make these responses appear to be acceptable. This format also meant that a 'yes' response could provide a negative answer, for example "Did you find the assessment boring". The questionnaire was printed on a bright light green paper so it stood out and could be differentiated from the therapist questionnaire. At the end of the interview the patient was invited to make additional comments.

Piloting the questionnaires

Bennet and Ritchie (1975) highlighted the need for researchers to "rigourously pre-test forms... to achieve the layout that reduces to a minimum errors by interviewer, respondents, coding clerks, or key punch operators" (p. 27). The researcher was to code and enter the data into the computer for analysis herself. Therefore, the focus of the pilot survey of these questionnaires was on questionnaire format and question clarity. The therapist survey was piloted by five occupational therapists working at two local hospitals. The therapists were asked to administer the SOTOF on one patient with a primary diagnosis of stroke, then complete the survey, time the length it took to complete the survey, and then discuss the whole procedure with the researcher in an unstructured interview. In the interview the therapists were invited to comment on all the materials (information letter, test manual and survey) with particular emphasis on ambiguity and any areas of difficulty. In addition to this verbal

feedback the researcher examined the therapists' questionnaire responses for signs of miscomprehension or unexpected response formats.

The patient survey was piloted on five participants who were in-patients at a local hospital. The patients were all 60 years old or over and had a primary diagnosis of stroke. Ethical permission for the study was provided by the Wandsworth Health Authority Ethical Committee, and subjects were recruited to the study via their consultant. The SOTOF was administered to the patient by the researcher (first author). Following test administration the patient was interviewed by a nurse or occupational therapist on the ward. The interviewer used the standardised patient questionnaire. On completion of the interview both the patient and interviewer discussed the whole procedure with the researcher. The questionnaire responses were reviewed for signs of miscomprehension or unexpected response formats.

No problems were identified with the content or administration procedures for either of the questionnaires and, therefore, no changes were made. Data regarding the length of time it had taken therapists to complete different aspects of the study was used to give potential subjects, for the main study, an indication of the time commitment they were required to make. This information was outlined in two recruitment letters (copies can be seen in Laver's 1994, PhD thesis, Appendix Seven and Appendix Eight).

Procedure

Although the SOTOF had been developed in 1990 it was decided that the study should be delayed until February 1991 owing to the Christmas holidays. A letter, requesting volunteers, was submitted to the editor of the British Journal of Occupational Therapy (BJOT) at the beginning of January 1991 (see in Laver's 1994, PhD thesis, Appendix Six). The study ran from mid-February to April. Volunteers replied to the BJOT letter by both telephone and post. Those who contacted by telephone were given further details of the nature and purpose of the study and were recruited during this conversation. Others were sent a letter (see in Laver's 1994, PhD thesis, Appendix Seven) and were invited to telephone the researcher to be recruited. Other therapists were recruited when they offered their assistance informally at a range of meetings and lectures which were attended by the researcher. Therapists were mainly drawn from the four Thames Regions. During the telephone call the researcher explained that ethical approval for the study had been obtained. Therapists were asked to check with their hospital that this ethical approval would extend to the inclusion of their patients in the study.

Once recruited, the therapist was sent a package containing an introductory letter, copy of SOTOF manual, a set of SOTOF assessment forms, a therapist questionnaire, a patient questionnaire, and a stamped, addressed envelope. The stamped, hand addressed envelopes were provided to encourage the return of questionnaires. The introductory letter (see in Laver's 1994, PhD thesis, Appendix Eight) explained the content of the package, criteria for patient subjects, the procedure to be followed, and an invitation to telephone the researcher if they had any queries. The therapist was provided with a copy of the SOTOF test manual and forms. They were instructed to administer the SOTOF to at least one patient. Two criteria were given for the selection of patients; they should be 60 years or over and

have a primary diagnosis of stroke. The SOTOF protocols and forms were designed to follow the same format for each of the four ADL tasks and therefore looked very similar. To differentiate the papers relating to each task the relevant protocols and observational checklist forms were colour coded. Therapists were given four weeks to complete the study. If therapists were unable to complete the research within four weeks they were requested to telephone the researcher. Non-respondents were followed up with a telephone call during the week after their deadline date.

Results

Analysis was undertaken using Clinstat software. Quantitative data was coded and then entered onto computer. Qualitative data was surveyed, themes which emerged from the qualitative responses to open ended questions were selected and labelled, and data was then coded and entered onto the computer. Clinstat was used to produce frequencies, percentages and means.

Description of samples

Two labels, to represent the two subject groups, will be used to clarify the presentation of results. The occupational therapist subjects will be referred to as "therapists" and the participants with stroke will be referred to as "patients". Demographic data for these samples was obtained in response to questions one and two on the therapist survey and the opening section of the patient survey.

Description of therapist sample

The therapist sample comprised of 104 subjects. The BJOT letter yielded 66 therapists and an additional 38 were recruited at meeting and lectures. Forty-four of the 104 therapists completed the study, giving a response rate of 42.3 percent which was in line with the expected 40-60 percent response rate (Bennett and Ritchie, 1975). Non-respondents fell into three main categories, lack of suitable patient subjects, illness and holidays. The majority of therapists were able to obtain ethical clearance for conducting the study on their patients based on the ethical approval obtained by the researcher from her own health authority; one non-respondent was unable to complete the study because ethical permission would have been required from his own health authority. The 44 therapists in the respondent sample comprised of all grades from basic grade to a head II occupational therapist (Table 5.1). They had been qualified as occupational therapists from less than one year to thirty years; length of experience with older people and people with a diagnosis of stroke were similarly distributed (1 - 30 years' experience with both patient groups). The therapists were drawn from a wide range of clinical specialties (Table 5.2).

Table 5.1: Therapists' Grades (n = 44)

Grade	Percentage of sample with that grade
Basic Grade	13.6% (n = 6)
Senior II	36.4% (n = 16)
Senior I	29.5% (n = 13)
Head IV	9.1% (n = 4)
Head III	2.3% (n = 1)
Head I	2.3% (n = 1)
Other: Occupational Therapist or Senior	6.8% (n = 3)

Table 5.2: Therapists Current Clinical Area (n = 44)

Clinical Area	Percentage of therapists working in that clinical area
Gerontology	38.6 % (n = 17)
Medical	15.9% (n = 7)
Gerontology - stroke patients	13.6% (n = 6)
Neurology	13.6% (n = 6)
Neurology - stroke patients	6.8% (n = 3)
Rehabilitation Unit	2.3% (n = 1)
Out Patients	2.3% (n = 1)
Surgery and Orthopaedics	2.3% (n = 1)
Medicine and Surgery	2.3% (n = 1)
Research	2.3% (n = 1)

Description of the patient sample who were administered the SOTOF

The therapists tested a total of 48 patients on SOTOF. Thirty-five therapists tested between one to three patients each, one therapist observed a colleague administering SOTOF to a patient. Eight therapists were unable to administer the test but examined the SOTOF manual and answered a proportion of the survey questions on this basis. Of the 48 patients tested, 20 were male and 28 were female. This slight bias towards female patients is reflective of the age-sex ratio of the general population for this age group. Participants' ages ranged from 62 to 89 years (mean 74.7, s.d. 6.92, missing data = 5). Sixteen participants had a left cerebral hemisphere lesion, 24 had a right cerebral hemisphere lesion, and 7 had a stroke of unknown origin (missing data = 1). Fifty-four point two percent of the subjects had a secondary diagnosis. A total of twenty six different secondary diagnoses were identified. The most frequent secondary diagnoses were: hypertension (n = 6); diabetes (n = 4); and osteoarthritis (n = 3).

Description of the patient sample that completed the Face Validity Survey

Of the 48 patients administer the SOTOF, 44 completed the face validity questionnaire, giving a response rate of 83.3 percent. Therapists reported that the four patients who did not complete the survey had expressive dysphasia. The patient

sample comprised 15 male and 25 female participants aged 62 to 87 years. Thirteen participants had a left cerebral hemisphere lesion, 20 had a right cerebral hemisphere lesion, and 3 had a stroke of unknown origin (missing data = 3).

Results of the occupational therapist survey: Content validity

Questions three and eighteen pertained to the content validity of SOTOF. These questions had an open format and produced qualitative data. Question three related to the therapists' perceptions of what SOTOF tests and eighteen to therapists' perceptions of the deficits which SOTOF could be used to highlight specifically in older people with a diagnosis of stroke. All the constructs listed for both questions were collated.

SOTOF is used to collect data on several different levels, independence in ADL, skills, performance components and neuropsychological deficits. Two different levels of data, performance components and neuropsychological deficits, were recorded on therapists' questionnaires. For example, some therapists referred to global performance components, such as perception, whilst others listed specific neuropsychological deficits, such an ideomotor apraxia. For comparison all data was converted to the higher level of performance components. The frequency with which each performance component was mentioned by therapists was calculated and expressed as a percentage. Components were then placed in order with the most frequently mentioned component first. Both the order and percentages of the components were then compared for the two questions. Results are shown in Table 5.3

Eight constructs emerged as the content base of SOTOF. These were: perceptual function; sensory function; motor function; cognitive function; language; performance of Activities of Daily Living (ADL); visual function; and auditory function. The three most frequently identified performance components were perceptual, sensory and motor functions. These three emerged as the most frequently mentioned component for both therapists' perceptions of what was tested by SOTOF and those deficits which could be highlighted by the test for older patients with stroke. A small number of responses could not be grouped under any of the eight performance component headings and were labelled "other factors".

Therapists were also asked to indicate what deficits, if any, they had highlighted in their patient. Again responses were given on several levels, performance components (motor function, sensory function), global deficit terms (recognition, comprehension) and specific neuropsychological deficits (ideomotor apraxia, agnosia, and visual neglect). Five of the 48 patients were reported to have had no deficits. All the other patients had been identified as having at least one deficit. Several therapists reported that SOTOF had been used to identify the following deficits: sensory deficit (n = 10); motor function (n = 10); spatial relationships (n = 6); tactile discrimination (n = 6); right/left discrimination (n = 5); neglect/inattention (n = 4); visual field loss (n = 3); body scheme (n = 3); and dressing apraxia (n = 3). Several additional deficits were identified in only one or two patients from the sample. The following additional deficits were identified for two patients in the sample: perseveration; figure ground discrimination; proprioception; abnormal tone; functional sitting balance; and expressive dysphasia. The following deficits were

identified for only one patient in the sample: recognition; agnosia; apraxia; ideomotor apraxia; motor apraxia; colour agnosia; stereognosis; comprehension; organisation; visual scanning; and speech deficit.

Table 5.3 Content Validity - Therapists perception of the content of SOTOF (n = 44). A Comparison of results to questions 3 and 18

Factor : area / deficit	Percentage factor occurred in responses to Question 3	Order of frequency factor occurred in response to Question 3	Percentage factor occurred in responses to Question 18	Order of frequency factor occurred in response to Question 18	Percentage difference in occurrence between questions 3 & 18
Perceptual	77.3%	1st	63.7%	1st	13.6%
Sensory	50%	2nd	27.3%	2nd	22.7%
Motor	36.7%	equal 3rd	18.2%	equal 3rd	18.5%
Cognitive	36.7%	equal 3rd	13.7%	5th	23%
Language	25%	4th	18.2%	equal 3rd	6.8%
Activities of Daily Living	22.8%	equal 5th	18.2%	equal 3rd	4.6%
Visual	22.8%	equal 5th	16%	4th	6.8%
Auditory	6.9%	7th	6.9%	7th	0
Other Factors	13.7%	6th	18.2%	equal 3rd	4.5%

Key to: Other Factors
Question 3: What areas do you feel this assessment tests? Other factors listed were: 'general screen'; 'as listed'; 'ability to perform task'.
Question 18: What deficits do you think this assessment could be used to identify in elderly stroke patients? Other factors listed were: 'as listed'; 'further assessment required'.

Clinical Utility

With regards to the **test manual and forms**, the majority of therapists (54.5%) indicated that the instructions were fairly easy to understand and to follow. A similar percentage (52.3%) also indicated that the protocols were fairly easy to follow and half of the therapists found the SOTOF forms easy to complete (Table 5.4).

Table 5.4: Therapists' perceptions of the SOTOF instructions, protocols and forms (n=44)
(Values = number of therapists indicating each descriptive rating)

Question: Were the... easy to...?	Impossible	Difficult	Fair	Easy	Very Easy
Were the instructions. easy to understand? (missing data = 0)	-	11.4% (n = 5)	54.5% (n = 24)	34.1% (n = 15)	-
Were the instructions easy to follow? (missing data = 0)	-	9.1% (n = 4)	52.3% (n= 23)	38.6% (n = 17)	-
Were the protocols easy to follow? (missing data = 1)	2.3% (n = 1)	9.1% (n = 4)	52.3% (n= 23)	34.1% (n = 15)	-
Were the forms easy to fill in? (missing data = 3)	-	6.8% (n = 3)	34.1% (n = 15)	50% (n = 22)	2.3% (n = 1)

With regards to **materials**, 72.7 percent found them easy to obtain and very high proportions of the sample reported that the materials were appropriate for their clients and easy to carry, clean and store (Table 5.5).

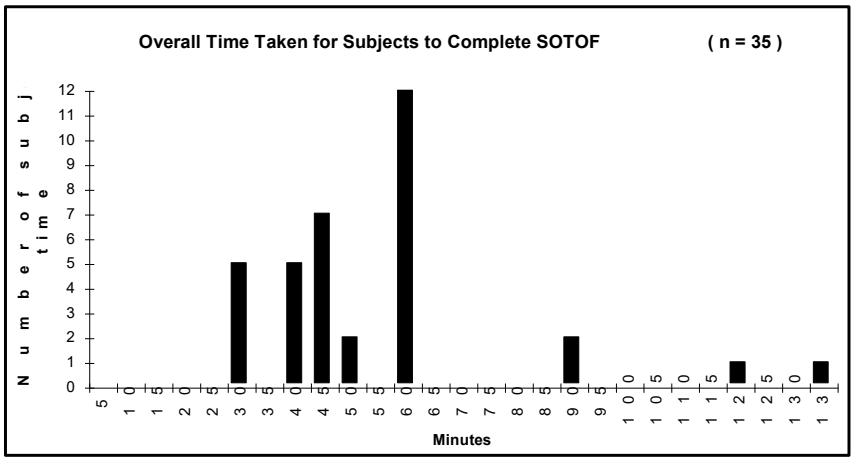
Table 5.5: Therapists' perceptions of SOTOF materials (n = 44)

Were the materials used..?	Yes	No	Missing Data
Easy to obtain	72.7% (n = 32)	22.7% (n = 10)	4.5% (n = 2)
Appropriate for your client	86.4% (n = 38)	6.8% (n = 3)	6.8% (n = 3)
Easy to carry	86.4% (n = 38)	4.5% (n = 2)	9.1% (n = 4)
Easy to clean	90.9% (n = 40)	2.3% (n = 1)	6.8% (n = 3)
Easy to store	88.6% (n = 39)	2.3% (n = 1)	9.1% (n = 4)

Administration time

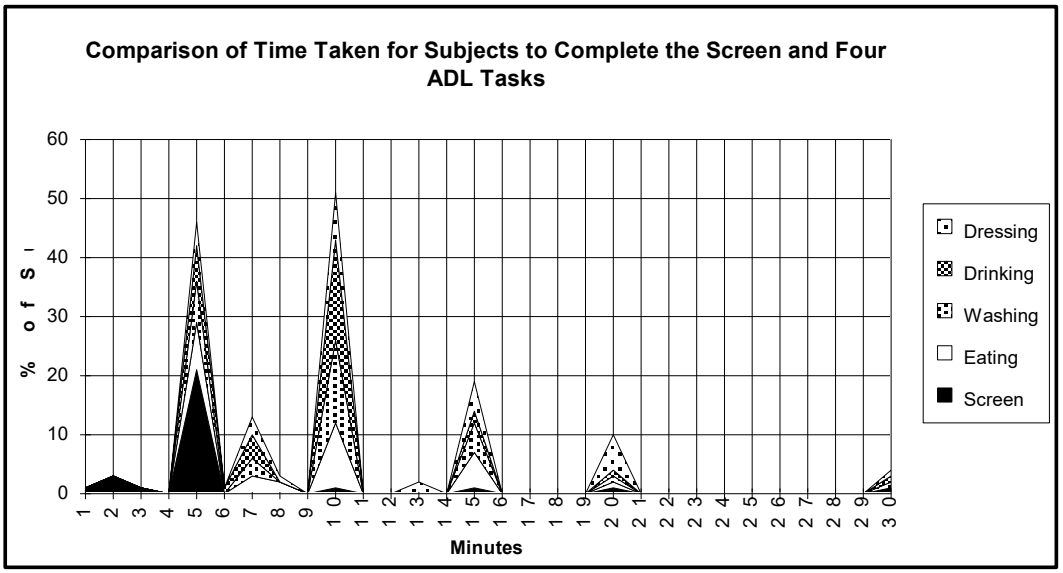
Therapists were asked to record data on the length of time taken to administer each sub-test and the whole SOTOF test. Times are illustrated in Figures 5.1 to 5.7. The majority of subjects took 60 minutes or less to complete the whole test (see Figure 5.1). Total time taken to administer all parts of SOTOF ranged from 30 minutes to 2 hours and 15 minutes, with only four subjects taking over an hour to complete the test.

Figure5.1: Graph showing Overall Time Taken by Subjects to complete SOTOF



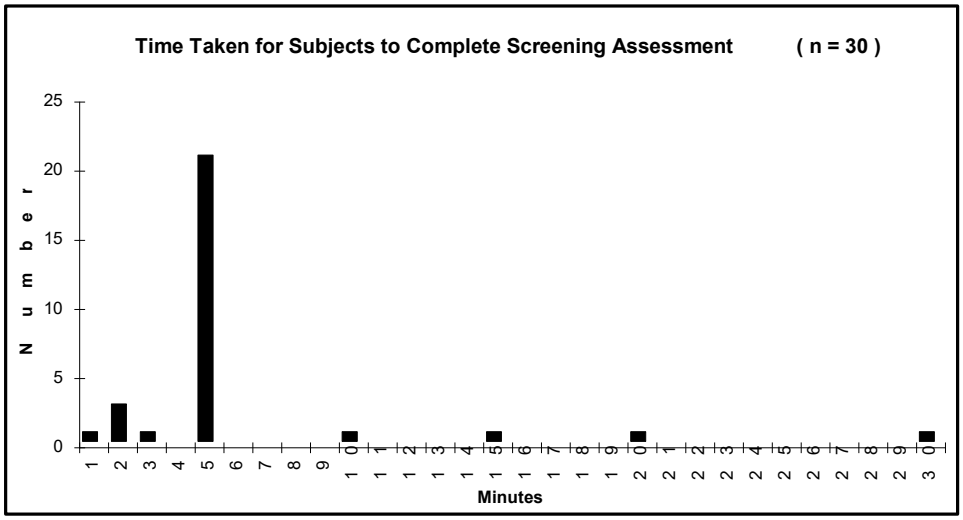
The five sub-tests (Screen, Eating, Washing, Drinking and Dressing Tasks) generally took between 5 to 20 minutes for the majority of subjects. Figure 5.2 provides a comparison of time taken for subjects across each of these five sub-tests.

Figure 5.2: Graph showing comparison of times taken across SOTOF sub-tests



The range of time taken to complete the screening test was 1 to 30 minutes. The majority of patients (n = 21) completed the screening test in 5 minutes (Figure 5.3).

Figure5.3: Graph showing the Time Taken on the Screening Test



Both the Eating and Washing tasks took between 5 to 15 minutes for most subjects to complete. The majority took 10 minutes to complete these tasks; 11 of the 31 subjects took 10 minutes to complete the Eating task (Figure 5.4) and 14 of the 31 subjects took 10 minutes to complete the Washing task (Figure5.5).

Figure 5.4: Graph showing the Time Taken by subjects to complete the Eating Task

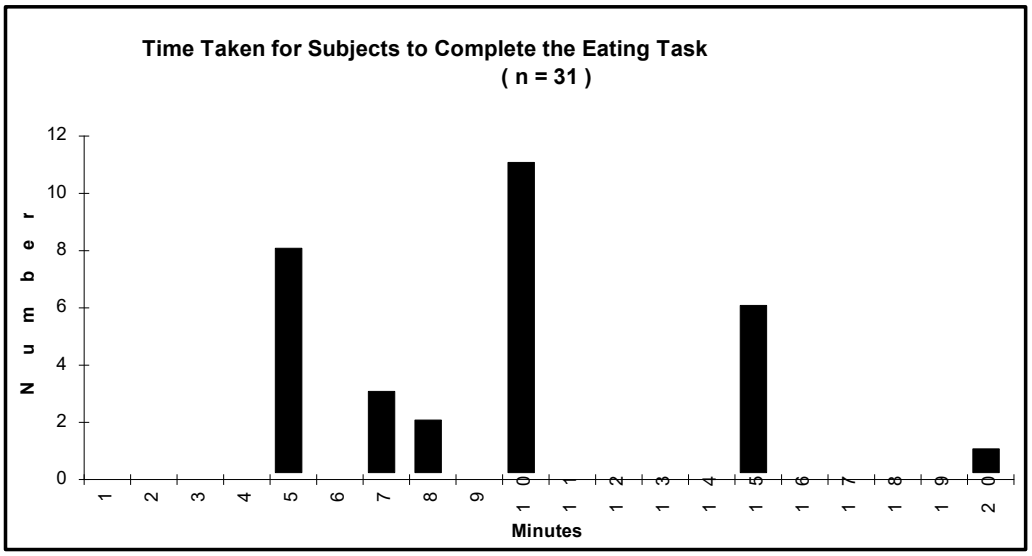
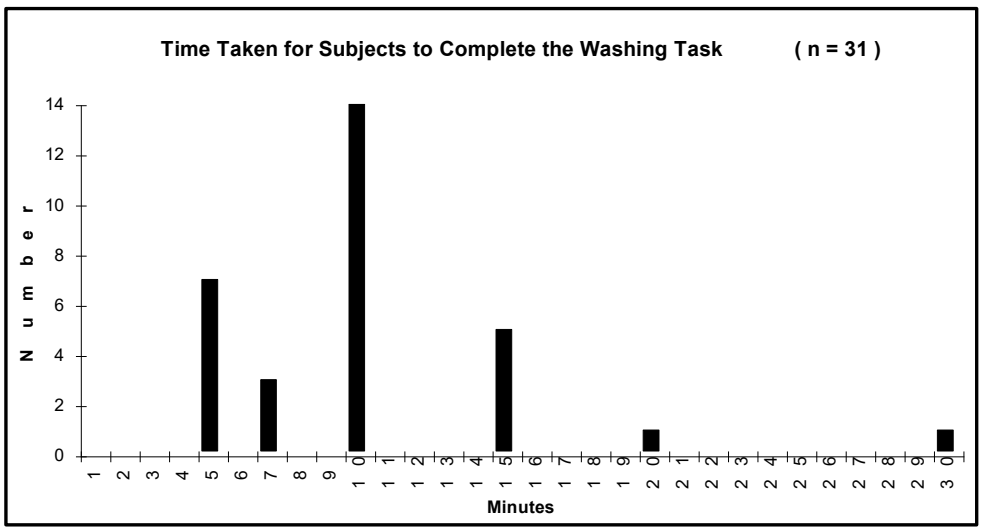
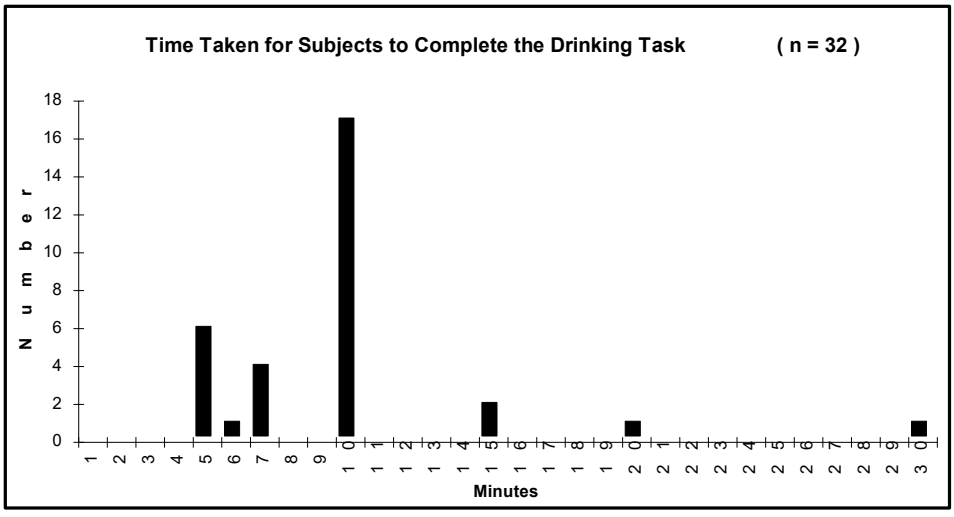


Figure 5.5: Graph showing the Time Taken by subjects to complete the Washing Task



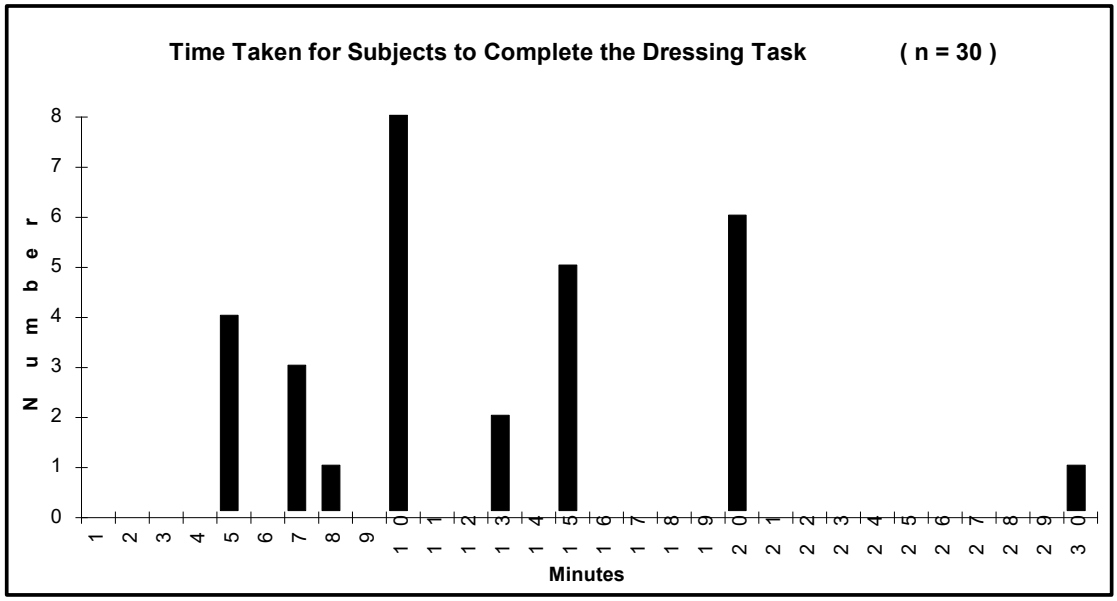
The Drinking task appeared to be slightly quicker and was completed by the majority in 5 to 10 minutes (see Figure 5.6). Seventeen of the 32 subjects complete the Drinking task in 10 minutes. Only four subjects took longer than 10 minutes to complete the task.

Figure 5.6: Graph Showing Time Taken by subjects to complete the Drinking Task



The Dressing task appeared to take little longer than the other four tasks and was usually completed in 5 to 20 minutes (see Figure 5.7). Four of the 30 subjects took only 5 minutes, 8 subjects took 10 minutes, 5 subjects took 15 minutes, and 6 subjects took twenty minutes to complete the Dressing task.

Figure 5.7: Graph showing the Time Taken by subjects to complete the Dressing Task



6 Enhancing the Dynamic Assessment Phase of the SOTOF (2nd edition)

Aim and objectives

The aim of this study was to undertake the first stage of a content validity test-development process to enhance SOTOF's dynamic assessment component. Content validity studies may draw upon 'literature relevant to the construct' and 'from other assessment instruments (i.e., borrowing items from other instruments that have demonstrated validity)' (Haynes, Richard & Kubany, 1995, p.247). This first stage study involved a review of literature and other occupational therapy dynamic assessments, for older people and / or for people with neurological conditions, to inform recommendations to improve SOTOF's dynamic assessment element. The objectives were to: develop an additional section in the SOTOF manual to explain how the dynamic element is assessed and reported; to develop further examples in the instruction cards to aid dynamic assessment related to each SOTOF test item; and to review and consider the scoring method.

Methodology

Three search strategies were developed to access different bodies of literature (see Tables 6.1, 6.2 and 6.3). The databases used were: the Cumulative Index of Nursing and Allied Health Literature (CINAHL); the Allied and Complimentary Medicine Database (AMED); OTSeeker; Medical Literature Analysis and Retrieval System Online (Medline); and Google Scholar. (Justification for the selection of these databases is presented in Table 6.4).

Search one: involved searching for literature that explained / used dynamic assessment within neurology (see Table 6.1).

Search two: involved the review of literature and test manuals of specific dynamic assessments. An initial search identified four dynamic assessment tools used by occupational therapists, developed for people with neurological conditions and suitable for the same populations as the SOTOF: the Contextual Memory Test (CMT; Toglia, 1993); the Executive Function Performance Test (EFPT; Baum & Wolf, 2013); the Learning Potential Assessment Device (LPAD; Feuerstein, Falik & Feuerstein, 1995); and the Dynamic Lowenstein Occupational Therapy Cognitive Assessment-for Geriatric Use (DLOTCA-G; Katz, Livni, Bar-Haim Erez & Averbuch, 2011). As the SOTOF was developed for older people (Laver & Powell, 1995), the DLOCTA-G was chosen, instead of DLOCTA (Katz et al., 2011). A search was conducted to identify any further literature related to these four dynamic assessment tools. This literature was appraised and then findings compared and contrasted with the SOTOF (see Table 6.2).

Search three: involved identifying literature related to how occupational therapists use dynamic assessment. Using all the search concepts provided only one result, however, when removing the first concept terms (see Table 6.3; dynamic concept /

model / theory / assessment) there were 1412 results, suggesting the term 'dynamic' was not widely used or adequately researched. Therefore, the search strategy included a snowball sampling strategy whereby relevant references in the reference lists of identified literature were sought. It also involved searching for the work of key authors related to dynamic assessments. In addition to journal articles, occupational therapy textbooks on grading, scaffolding and cueing were identified from search three.

The following inclusion and exclusion criteria were used for *search one*:

Inclusion criteria: 1) Literature published within fifteen years, i.e. from 2001 (as this was predominantly when the dynamic assessment concept appears to have evolved); 2) Primary research studies should have ethical approval; 3) Literature relevant to neurological impairment / function; 4) Designed to be used by health professionals.

Exclusion criteria: 1) Literature related to children or infants; 2) Dynamic assessment in educational literature. For *search three*, the 3rd inclusion criteria was removed and 4th changed to 'designed to be used with occupational therapists'.

Ethics approval was not required for this secondary research study. The ethics checklist for York St John University was completed and no ethical considerations were identified.

Table 6.1: Search strategy 1

		OR	**OR**	**OR**	**OR**
Concept 1 Search Terms		Dynamic assessment	Dynamic model	Dynamic concept	Dynamic theor*
Concept 2 Search Terms	AND	Neurology	Neurological impairment	Neurological function*	

Note: * is used for truncation and is used to search for terms beginning with the same letters.

Data analysis

Data analysis was undertaken using a narrative approach (Aveyard &Sharp, 2013). The findings were then considered and applied in terms of their relevance to the enhancement of SOTOF's dynamic assessment component. The data analysis was undertaken by the first author and then discussed with, and reviewed by, the second author. After drafting new materials for the SOTOF the proposed test materials were reviewed by an occupational therapy academic with expertise in dynamic assessment and some amendments were made based on his feedback.

Table 6.2 Search strategy 2

		OR	**OR**	**OR**	**OR**	**OR**	**OR**	**OR**	**OR**
Concept 1 Search Terms	AND	Dynamic assessment	Dynamic model	Dynamic concept	Dynamic theor*				
Concept 2 Search Terms	AND	Adult	Older person	Older people	Elderly				
Concept 3 Search Terms	AND	Executive function performance test	Contextual memory test	The application of cognitive functions scale	Dynamic Lowenstein occupational therapy cognitive assessment – for geriatric use	EFPT	CMT	ACFS	DLOCTA-G

Table 6.3: Search strategy 3

		OR	**OR**	**OR**	**OR**
Concept 1 Search Terms		Dynamic assessment	Dynamic model	Dynamic concept	Dynamic theor*
Concept 2 Search Terms	AND	Occupational therap*			
Concept 3 Search Terms	AND	grading	scaffolding	prompting	cueing

Table 6.4: Justification for databases chosen

Database	What it offers	Why was it included in the study?
CINAHL	Journals and publications of high quality nursing and allied health professionals' literature. Covers health sciences and seventeen allied health professions. Has advance search options and allows the researcher to view cited references (EBSCO Health, 2015).	Allows access to a variety of sources, not journals alone, including health care books, dissertations and conferences. Provides full-text material for the majority of sources.
AMED	More of a focus on alternative treatments but is designed for therapists, physicians and medical researchers. Hundreds of journals can be accessed via AMED. It is developed by a company who are known for their use of scientific, technical and medical information (EBSCO Host, 2015).	Key subjects within AMED include occupational therapy and rehabilitation. It is designed for therapists alongside medical researchers, therefore, this database was used within the study to support a more scientific and technical view.
OTSeeker	OTSeeker provides easy access and encompasses a variety of resources specifically relevant to occupational therapy practice. 1,000 different journals have been used within OTSeeker; these have been sought from bibliographic databases (Bennett et al, 2003).	A specialised database for occupational therapy practice. To enable occupational therapists to work using the evidence-base it is vital to integrate clinical experience and client perspectives with the research literature (Bennet et al, 2003).
Medline	Database of life sciences and biomedical information. Journals included from a wide range of health professions. Many of these are journal articles in life sciences (U.S. National Library of Medicine, 2015).	Provides a substantial literature from a variety of professions useful for gaining a wider view.
Google Scholar	Provides a freely accessible search engine with a wide variety of sources including academic publishers (Slater, 2013).	Provides wide range of literature. It allows a researcher to also search a specific article and sometimes it is more accessible than on some other databases.

Table 6.5: Types of literature included and reviewed

Reference	Test Manual	Webpage	Primary research articles (not RCTs)	Randomised control trial (RCT)	Book or book chapter	Journal articles, including secondary research
Arbesman M, Lieberman D and Metzler CA (2014)						✓
Baum CM and Wolf TM (2013)	✓					
Cotrus A and Stanciu C (2014)						✓
Doucet BM (2012)						✓
Falik LH and Feuerstein R (2005)						✓
Feuerstein R, Falik LH and Feuerstein R (1995)	✓					
Feuerstein R, Falik LH and Feuerstein R (2003)					✓	
Feuerstein R, Falik L, Rand Y and Feuerstein RS (2002)					✓	
Golledge J (2006)					✓	

Hadas-Lidor N and Weiss P (2005)					✓	
Hadas-Lidor N, Weiss P and Kozulin A (2011)					✓	
Haywood CH and Lidz CS (2007)					✓	
Haywood HC and Miller MB (2003)				✓		
Hersch GI, Lamport NK and Coffey MS (2005)					✓	
Hessels-Schlatter C (2002)			✓			
Katz N, Livni L, Bar-Haim Erez A and Averbuch S (2011)	✓					
Katz N, Erez ABH, Livni L and Averbuch S (2012)				✓		
Katz N, Averbuch S and Bar-Haim Erez A (2012)				✓		
Koenigsknecht C and Smith H (2012)		✓				
Laver AJ and Powell GE (1995)	✓			✓		

Results

Literature was identified from journal articles (primary research and secondary studies), text books, test manuals and web based resources. Table 6.5 presents the 34 pieces of literature retrieved and reviewed, categorised by type.

Dynamic assessment applied to neurological assessment.

Dynamic assessments are suggested to be advantageous over the use of static assessments within neurology owing to their ability to evaluate and enable the therapist to understand abilities, learning potential and possible changes that may occur as a result of guidance, specifically when predicting outcomes for service users and evaluating recovering functions (Hadas-Lidor, Weiss & Kozulin, 2011: Katz & Bar Haim Erez et al., 2012). Dynamic assessments can reduce cultural and educational bias, as people can learn, adapt and receive feedback throughout and clinicians can understand how the person learns best and what level of support is required (Uprichard, Kupshik, Pine & Fletcher, 2009). The generation of useful information from dynamic assessments has been evidenced, for example in studies related to adults with schizophrenia and older people with dementia (Wiedl, Schottke, Garcia & Dolores, 2001). Rather than focusing on normative data and typical performance, dynamic assessments allow clinicians to: focus on individual variations, changes and barriers to performance: and explore how individuals can improve their performance with some form of guidance (Cotrus & Stanciu, 2014; Toglia & Cermak, 2009). Dynamic assessments can provide occupational therapists with information to guide intervention planning and realistic goal setting (Katz & Bar Haim Erez et al., 2012; Toglia, 2011). Clinicians need to think differently about assessments and understand that abilities are changeable and sensitive to guidance (Toglia & Cermak, 2009). Dynamic assessments are useful within group assessments; however, in comparison with individual assessments, group assessments may generate less rich data (Haywood & Miller, 2003). Haywood and Lidz (2007) mentioned that dynamic assessments can be too time consuming for everyday practice, however, they still advocated for their use.

Dynamic assessment in relation to cognition

The ability to transfer information and learning are critical areas of cognition and can be directly addressed through dynamic assessments (Toglia, 2005); this is important for people with neurological conditions who often need to relearn skills. The interactive element of dynamic assessments provides an opportunity to maximise engagement and motivation (Cotrus & Stanciu, 2014; Toglia & Cermak, 2009). Dynamic assessments use cues, task alterations, physical context cues, self-awareness strategies and guidance assistance to identify a person's cognitive modifiability (Toglia, 2011). Dynamic assessments view cognition as modifiable, they provide a direct link to intervention and are flexible and person-centred (Toglia, 2011). Further research of dynamic assessments in regards to neurology and cognition is needed (Toglia, 2011).

Review of four dynamic assessment tools

The **EFPT** (Baum & Wolf, 2013) uses a standardised, six point (0-5), graduated cueing system to identify what the person can do and how much assistance is required. For example: if a person is independent completing a task he / she scores zero; if the clinician completed the task for the person he / she scores five. These scores are totalled at the end of each task. EFPT consists of pre-test questions and four tasks examining initiation, executive functioning and completion of the task. The cueing system is described with specific instructions for each cueing level (Baum & Wolf, 2013).

The **CMT** (Toglia, 1993) was designed to assess memory strategy use, recall and recognition and awareness of memory capacity (metacognition) for adults with memory problems. The first phase is standardised; and the second phase is dynamic (Toglia, 2011). Within phase two, cues can be used during the delayed recall task after the use of questioning strategies. Cues should be provided on two levels: a more general cue; and then more specific. The manual provides clear guidance for clinicians, presenting examples of cues that could be given for both levels of prompting for each task (Toglia, 1993). A summed score is provided for cued recall and recorded on the 'summary of findings' sheet.

The **LPAD** (Feuerstein, Falik & Feuerstein, 1995) can be used for people of all ages; it uses the mediation method (Feuerstein, Falik, Rand & Feuerstein, 2002) and is based on the theory of cognitive modifiability, which considers cognition not to be fixed (Feuerstein Institute, n.d.). The LPAD requires training and is suitable for occupational therapists to use (Feuerstein Institute, n.d.; Missiuna, 1987). It focuses on isolating three levels (input, elaboration and output) to understand how the phase(s) intervene on performance (Feuerstein, Falik & Feuerstein, 2003). There are sixteen LPAD sub-tests, focusing on: visual motor skills; perceptual organisation motor skills; perceptual organisational skills; memory; higher cognitive processes; and mental operations (Falik & Feuerstein, 2005). The assessor is able to detect conditions that are preventing the person from higher levels of performance to support intervention planning (Feuerstein et al., 2003; Falik & Feuerstein, 2005). A limitation of the LPAD, in terms of its relevance to the SOTOF, is that it uses some abstract psychological tests that have limited 'real-world' functionality.

The **DLOCTA-G** (Katz et al., 2011) was designed for use with adults over the age of seventy (Smith, 2012). It consists of 24 sub-tests focusing on eight cognitive areas: orientation; awareness; visual perception; spatial perception; praxis; visuomotor construction; thinking operations; and memory. The DLOCTA-G uses a graduated prompting schedule providing the dynamic element (Smith, 2012). This gives a structured and systematic approach to support clinicians in understanding what the person requires to master the task (Katz & Bar Haim Erez et al., 2012). There are five levels of mediation; level one is more general through to level five involving copying or subtracting the amount of stimuli (Katz et al., 2011). The DLOCTA-G was found to be useful in the evaluation of clients who have had a stroke (Katz & Bar Haim Erez et al., 2012). The dynamic element is scored on an ordinal scale (Smith, 2012). Like the LPAD, a limitation of the DLOCTA-G, in terms of its relevance to the SOTOF, is its lack of focus on everyday tasks.

How occupational therapists use dynamic assessment

Occupational therapists have specific skills and knowledge making them well equipped to undertake dynamic assessments (Toglia, 2005) and understand participation in the environment and how it can be adapted to enable performance through the use of cues, prompts and stimuli (Golledge, 2006; Hadas-Lidor & Weiss, 2005; Liddle & Eagles, 2014). Occupational therapists have understanding of self-awareness, processing strategies, cues, grading, scaffolding and how they are modifiable and vital when understanding occupational performance (Toglia, 2005) and have specialist knowledge regarding the use of clinical reasoning, therapeutic use of self and activity analysis (Arbesman, Lieberman and Metzler, 2014). Several forms of clinical reasoning are utilised in dynamic assessment and, in particular: abductive reasoning may be applied in the initial generation of hypotheses based on therapists' observations of the person's behaviour / responses; and a hypothetico-deductive approach underpins diagnostic reasoning (Rogers & Holm, 1991). The use of activity analysis allows therapists to understand the physical and cognitive requirements to complete a task, enabling them to identify skills and the context required in line with the service users' abilities (Hersch, Lamport & Coffey, 2005; Missiuna, 1987). Missiuna (1987) was one of the first occupational therapy authors to identify and discuss the relevance of dynamic assessment for occupational therapy. She discussed how occupational therapists are uniquely equipped to work within the dynamic concept. Dynamic assessments can be applied to other assessments, for example, observing a client struggle and intervening at a suitable level using clinical reasoning, with 'the exact level of assistance which the client would need in order to be independent …determined in a systematic manner through the graduated provision of prompts or aids' (Missiuna, 1987, p.18-19). When undertaking dynamic assessment therapists need an interactive and motivational communication style, contrary to the more usual 'neutral' stance taken therapists during standardised assessment; 'this type of intervention is termed 'mediation' in that the therapist mediates between the client and the task' (p. 19).

For rehabilitation, therapists need to be aware of techniques, approaches and modifications required for a person to master a task. Dynamic assessments can be used to provide this information during the assessment phase, otherwise such techniques / approaches / modifications are identified during the intervention phase through a more lengthy process of trial and error (Missiuna, 1987). Nott, Chapparo and Heard (2008) found that occupational therapists' use of dynamic assessments improved participants' ability to use strategies during occupations in comparison to other interventions. Toglia (1998) developed a structured grading system, providing clinicians with a guided, more effective way of distributing cues. Sanderson and Gitsham (1991) discussed the Interactive Sequence comprising four graduated levels of independence / dependence, including 'preference', the level at which a person chooses to initiate an activity, and the 'co-active level' where the therapist and client are working as one.

Discussion

Two issues emerged during the analysis of obtained literature. Firstly, some authors do not specifically use the term 'dynamic'. For instance, Golledge (2006) discussed prompting, verbal guidance and cues, all components of dynamic assessment, but

did not use the word ‘dynamic’. Secondly, the terms ‘prompt’ and ‘cue’ are used interchangeably. The authors reflected on the use of the terms ‘cue’ and ‘prompt’ within SOTOF and conducted a review of literature to explore definitions of these terms. The definitions provided by the Texas Education Agency (2011) best explained the differences between cues and prompts: ‘Cueing (general assistance) is an action intended to encourage a [person] to initiate or continue a task that he or she had previously executed; a cue is a hint or nudge in the right direct that does not provide a direct answer’ (p.1). Whereas ‘Prompting (specific assistance) is an action intended to directly assist a [person] with the completion of a task; a prompt pulls the [person] through each step to the end of the task and directly leads to the answer’ (p.1). Literature discussing how occupational therapists have the necessary skills to use dynamic assessments was identified. A dynamic assessment approach is relevant for occupational therapy practice (Missiuna, 1987) but further research is needed to evaluate occupational therapists’ use of dynamic assessment (Doucet, 2012). The findings (Hadas-Lidor et al., 2011; Katz & Bar Haim Erez et al., 2012; Toglia, 2011; Uprichard et al., 2009) suggested that dynamic assessments generate valuable information and provide occupational therapists with the tools to develop person-centred intervention plans. The dynamic assessments evaluated, (EFPT, CMT, LPAD and DLOCTA-G) provided different approaches to using dynamic assessment. However, all aim to generate knowledge on how the person best learns and how some form of guidance can increase his / her level of performance. Both the EFPT (Baum & Wolf, 2013) and the DLOCTA-G (Katz et al., 2011) use graduated cueing / prompting methods, the CMT (Toglia, 1993) focuses on strategy use and the LPAD (Feuerstein et al., 1995) applies a mediation approach

Application of findings to the SOTOF

After evaluating the four assessment tools, reviewing literature related to dynamic assessment and comparing findings to the SOTOF, it was clear there was potential to strengthen and formalise the SOTOF’s dynamic assessment component. Like the CMT (Toglia, 2011), it was decided that the first phase of SOTOF administration and the dichotomous ‘able’ or ‘unable’ scoring for each test tem would remain standardised. However, the second phase would be dynamic and would be applied whenever a person is unable to successfully undertake a test item. It was decided to develop a ‘graduated mediation protocol’ to be applied to all test items across the four SOTOF ADL tasks. The term ‘mediation’ was chosen, because therapists can mediate between the client and the task (Missiuna, 1987) using a number of different strategies, including cueing, prompting, assisting, modifying and demonstrating. The SOTOF six level graduated mediation protocol (see Table 2.2, earlier in Chapter 2) was developed by predominately applying the EFPT’s graduated cueing instructions (Baum & Wolf, 2013), the DLOCTA-G’s graduated prompting schedule (Katz et al., 2011) and the principle of co-active assistance (Sanderson & Gitsham, 1991)..Updated instruction cards were developed, providing examples of prompts / cues / modifications / assistance for each test item on the four ADL tasks for levels 1 – 4 of the six point SOTOF graduated mediation method. The beginning of the Screening Assessment was provided earlier in Figure 2.1 in Chapter 2.

The SOTOF record form was updated to incorporate scoring for the 0-5 level graduated mediation protocol (see Figure 2.2: Example of first part of the Screening Assessment Record Form, in Chapter 2). This scoring system was influenced by the

method used to score the EFPT (Baum et al., 2008). The item scores are totalled for each SOTOF ADL tasks, and the four task totals are then summed to produce an overall score (the higher the score, the more guidance / assistance required). An additional section was added to the summary parts of the record forms to prompt therapists to comment on the persons' learning potential, a key element of a dynamic assessment approach, and provide information on the most effective form of mediation to support future assessments and interventions. The end of the neuropsychological checklist has been updated in line with the introduction of the six level graduated mediation protocol (see Table 2.5, in Chapter 2). An additional section for the SOTOF manual was written to provide guidance on using the graduated mediation protocol, the revised record form and the neuropsychological checklist and summary form (see Figure 2.3).

Gaining feedback on the new SOTOF developments resulted in discussion regarding the equivalence of different forms of mediation at level 4 (see Table 2.2). However, it was decided that separating this level, into four levels (physical assistance, co-active assistance, modification, demonstration), could introduce confusion and not all potential expanded levels could be applied to the majority of SOTOF task items. For instance, if copying an action was separated in the hierarchy, this level would have to be omitted for some task items (e.g., when the person is asked to close their eyes to identify an object through touch). Level four should, therefore, be seen as an umbrella level; it maintains the hierarchical element, as it increases the input from the clinician from level three, but this can be done in a number of ways, depending on the test item and needs of the person (see level 4 in Table 2.2). To ensure the type of mediation is clear, the type of mediation used should be specified by the clinician on the record form; this instruction was also added to the manual to ensure clarity.

Relevance for occupational therapy practice and research

This study has provided evidence to suggest that occupational therapists use dynamic assessment in clinical practice. However, the phrase 'dynamic assessment' is not commonly used and maybe under-recognised amongst occupational therapists. There is little primary research evaluating occupational therapists' use of dynamic assessment in clinical practice. The findings were used to developed new materials to enhance and formalise SOTOF's dynamic assessment element. It is hoped this will provide occupational therapists with an updated, dynamic assessment tool for clients with suspected neurological impairment. Further research is required to examine the psychometric properties of the revised SOTOF, including the second stage content validity evaluation by an expert panel and evaluation of inter-rater and test re-test reliability for the proposed scoring method. A clinical utility study evaluating the application of the revised SOTOF in clinical practice would be beneficial.

Strengths and Limitations

Owing to cost restrictions the authors were unable to purchase the LPAD and DLOTCA-G test manuals; however, related literature was sought that discussed their administration, and explained how dynamic assessment is used within these tests.

There was limited literature identified from search three, so a snowballing strategy, involving referenced literature and authors in the field, was applied; this may reduce the replicability of the study. However, the use of several well-established databases and three separate searches located useful and relevant literature to inform the content validity of the proposed SOTOF graduated mediation protocol, related examples and scoring method. The use of two researchers to review the analysis and consider the application of findings to SOTOF, in addition to obtaining feedback from an academic, with a particular interest in dynamic assessments, was undertaken to increase trustworthiness and decrease bias.

Conclusion

Literature reviewed indicated the value of dynamic assessments for occupational therapy practice and identified that occupational therapists have the skills and knowledge required to undertake dynamic assessment. Elements of dynamic assessment (e.g. prompting, cueing, grading) are embedded into occupational therapists' practice, even where the term 'dynamic assessment' is not explicitly used. Dynamic assessments have been shown to provide therapists with in-depth knowledge and understanding of an individual's abilities and identify their learning potential and strategies for intervention. This study led to further development of SOTOF's dynamic assessment component, drawing on the appraisal of four dynamic assessment tools and review of literature related to dynamic assessment.

Recommended Future research

This study contributed to the first stage content validity for the SOTOF. The SOTOF 2nd edition now comprises a formalised dynamic assessment component using a six-level graduated mediation protocol and rating scale. SOTOF phase one administration remains standardised and the dynamic assessment (phase two) is applied to SOTOF test items the person is unable to do. A stage 2 content validity study involving review by an occupational therapy expert panel and a clinical utility study are now planned. Further studies to evaluate the inter-rater and test-retest reliability of the new 6 point ordinal scoring system and summed scores for the four ADL tasks is now required.

Appendix: Normative Standards

Normative standards are for time taken to perform SOTOF 1st edition (i.e. without the inclusion of the graduated mediation protocol)

***Table A.1**: Performance of normal sample: time taken on sub-tests for SOTOF (1st ed)*

Sub-test	*Time range (minutes and seconds)*	*Mean*	*Standard deviation (sd)*
Screening Assessment	29" – 3'54"	59.96	29.98
Task 1	1'31" – 7'12"	188.97	78.55
Task 2	2'5" – 7'8"	204.35	59.39
Task 3	1'3" – 6'	143.90	53.90
Task 4	1'3" – 4'40"	108.20	41.79
SOTOF (total time for all sub-tests)	7'16" – 23'9"	693.63 (11'33")	222.17 (3'42")

***Table A.2**: Time taken by normal sample on SOTOF Screening Assessment (1st ed) (n=64)*

Time ranges (minutes and seconds)	*Time grades*	*% of normal population obtaining time grade*
< 0'29"	Superior	-
0'29" – 0'36"	Above average	15.6% (n=10)
0'37" – 1'16"	Within normal limits	68.8% (n=44)
1'17" – 2'16"	Below average	14.1% (n=9)
2'17" +	Impaired	1.6% (n=1)

***Table A.3**: Time taken by normal sample on Task 1 (n-67): eating from a bowl using a spoon*

Time ranges (minutes and seconds)	*Time grades*	*% of normal population obtaining time grade*
< 1'35"	Superior	1.5% (n=1)
1'35" – 1'59"	Above average	14.9%% (n=10)
2'00" – 4'28"	Within normal limits	68.7% (n=46)
4'29" – 7'02"	Below average	13.4% (n=9)
7'03" +	Impaired	1.5% (n=1)

Table A.4: *Time taken by normal sample on SOTOF 1st ed. Task 2 (n=68): washing hands in a bowl*

Time ranges (minutes and seconds)	*Time grades*	*% of normal population obtaining time grade*
< 2'07"	Superior	1.5% (n=1)
2'07" – 2'36"	Above average	14.7% (n=10)
2'37" – 4'35"	Within normal limits	69.1% (n=47)
4'36" – 6'43"	Below average	13.2% (n=9)
6'44" +	Impaired	1.5% (n=1)

Table A.5: *Time taken by normal sample on SOTOF 1st ed. Task 3 (n=70): pouring and drinking*

Time ranges (minutes and seconds)	*Time grades*	*% of normal population obtaining time grade*
< 1'06"	Superior	1.4% (n=1)
1'06" – 1'40"	Above average	14.3% (n=10)
1'41" – 2'02"	Within normal limits	68.6% (n=48)
2'03" – 5'39"	Below average	14.3% (n=10)
5'40" +	Impaired	1.4% (n=1)

Table A.6*: Time taken by normal sample on SOTOF 1st ed. Task 4 (n=64): putting on a shirt*

Time ranges (minutes and seconds)	*Time grades*	*% of normal population obtaining time grade*
< 1'03"	Superior	-
1'03" – 1'14"	Above average	15.6% (n=10)
1'15" – 2'27"	Within normal limits	68.8% (n=44)
2'28" – 4'35"	Below average	14.1% (n=9)
4'36" +	Impaired	1.6% (n=1)

***Table A.7**: Total time taken by normal sample on SOTOF 1st ed. (n=41): Screening Assessment and four tasks*

Time ranges (minutes and seconds)	*Time grades*	*% of normal population obtaining time grade*
< 8'51"	Superior/above average	14.6% (n=6)
8'51" – 15'26"	Within normal limits	70.7% (n=29)
15'27" +	Below average/impaired	14.6% (n=6)

Normative standards for descriptive responses

***Table A.8**: Descriptions of use of objects given by normal elderly clients for Task 1 (n=86)*

Descriptive category	*% first description given (n=65)*	*% second description given (n=29)*
Eating/to eat with	27.9% (n=24)	5.8% (n=5)
For cereal	20.9% (n=18)	4.7% (n=4)
For fruit	3.5% (n=3)	8.1% (n=7)
For soup	7.0% (n=6)	5.8% (n=5)
For pudding/sweet	4.7% (n=4)	11.6% (n=10)
For food	-	2.3% (n=2)
Other descriptions	11.6% (n=10)	4.7% (n=4)

***Table A. 9**: Descriptions of use of objects given by normal elderly clients for Task 2 (n=86)*

Descriptive category	*% first description given (n=59)*	*% second description given (n=16)*
Washing/to wash	46.5% (n=40)	1.2% (n=1)
To wash my/your hands	5.8% (n=5)	2.3% (n=2)
To wash myself/yourself	7.0% (n=6)	1.2% (n=1)
Put water in the bowl	2.3% (n=2)	-
Drying/dry with the towel	-	8.1% (n=7)
Put soap on a flannel	-	1.2% (n=1)
Other descriptions	7.0% (n=6)	4.7% (n=4)

***Table A. 10**: Descriptions of use of objects given by normal elderly clients for Task 3 (n=86)*

Descriptive category	*% first description given (n=68)*	*% second description given (n=47)*
Drinking/to drink	25.6% (n=22)	25.6% (n=22)
Pouring/to pour	19.8% (n=17)	3.5% (n=3)
For water	9.3% (n=8)	5.8% (n=5)
For other named drink, e.g. milk	7.0% (n=6)	14% (n=12)
For serving drink	11.6% (n=10)	2.3% (n=2)
Other descriptions	5.8% (n=5)	3.5% (n=3)

***Table A.II:** Descriptions of use of objects given by normal elderly clients for Task 4 (n=86)*

Descriptive category	*% first description given (n=82)*	*% second description given (n=41)*
For warmth/to keep warm	32.6% (n=28)	12.8% (n=11)
To keep cold out	2.3% (n=2)	2.3% (n=2)
To wear/wearing	18.6% (n=16)	8.1% (n=7)
For doing up	8.1% (n=7)	8.1% (n=7)
To fasten/fastening	4.7% (n=4)	2.3% (n=2)
Sew button on	15.1% (n=3)	11.6% (n=10)
Other descriptions	2.3% (n=2)	2.3% (n=2)

***Table A.I2**: Descriptions of taste of food given by normal elderly clients for Task 1 (n=86)*

Descriptive category	*% first description given (n=49)*	*% second description given (n=19)*
Named actual food	38.4% (n=33)	15.1% (n=13)
Sweet	9.3% (n=8)	1.2% (n=1)
Sharp	2.3% (n=2)	1.2% (n=1)
Other descriptions	7.0% (n=6)	4.7% (n=4)

***Table A.13**: Descriptions of taste of drink given by normal elderly clients for Task 3 (n=86)*

Descriptive category	*% description given (n=28)*
Named actual drink	36.7% (n=23)
Tasteless (water)	2.3% (n=2)
Sweet	1.2% (n=1)
Other descriptions	2.3% (n=2)

***Table A. 14**: Descriptions of temperature of water given by normal elderly clients for Task 2 (n=86)*

Descriptive category	*% description given (n=40)*
Warm	23.3% (n=20)
Tepid	8.1% (n=7)
Luke warm	8.1% (n=7)
Other descriptions	7.0% (n=6)

Enlarged Written Instructions

To use these instructions, cover the page so that the only instruction visible is the one that you want the client to read. Otherwise print and laminate each separate instruction onto an individual card.

Screening Assessment

What is your name?

What is this?

Which is the pen?

Which is the cup?

Which is the spoon?

Copy what I do.

Hold this.

Now hold the cup in your other hand.

Which hand did you use to write with?

Which hand do you use to write with?

Do you use glasses, a hearing aid or any other equipment?

Task 1: Eating from a Bowl using a Spoon

I am going to ask you some questions and ask you to do something familiar. I want to see what you can do and what problems you may have, so that we can plan your treatment. Don't worry if you feel you cannot manage everything.

Please close your eyes. I am putting an object in your hand, and I want you to tell me what it is without looking.

What can you see on the table?

Which is the bowl?

Which is the mat?

Which is the spoon?

Put the spoon on the right of the bowl.

What do you use these objects for?

Without touching the objects show me how you would use them.

Mime how you would use these objects.

Touching the objects, show me how you would use these objects.

Touching the objects, show me how you would use them.

What colour is the bowl?

What colour is the spoon?

What colour is the mat?

Eat the food in the bowl using the spoon.

Can you taste the food?

How does it taste?

Have you finished all the food?

Put the spoon on the table on the left of the bowl.

Is that your left?

Put the spoon in front of the bowl,

Put the spoon behind the bowl,

Put the spoon in the bowl.

Task 2: Washing Hands from a Bowl

I am going to ask you some questions and ask you to do something familiar. I want to see what you can do and what problems you may have, so that we can plan your treatment.

Don't worry if you feel you cannot manage everything.

Please close your eyes. I am putting an object in your hand, and I want you to tell me what it is without looking.

What can you see on the table?

Which is the bowl?

Which is the soap?

Which is the towel?

Put the soap on the table on the right of the bowl.

What do you use these objects for?

Without touching the objects show me how you would use them.

Mime how you would use these objects.

Touching the objects, show me how you would use these objects.

What colour is the bowl?

What colour is the soap?

What colour is the towel?

Wash your hands in the bowl using the soap, and dry your hands on the towel.

Lift your hand into the bowl of water with your other hand.

What temperature is the water?

Dry your hands on the towel.

Put the soap on the table on the left of the bowl.

Is that your left?

Put the soap in front of the bowl,

Put the soap behind the bowl.

Put the soap in the bowl

Task 3: Pouring and Drinking

I am going to ask you some questions and ask you to do something familiar. I want to see what you can do and what problems you may

have, so that we can plan your treatment. Don't worry if you feel you cannot manage everything.

Please close your eyes. I am putting an object in your hand, and I want you to tell me what it is without looking.

What can you see on the table?

Which is the jug?

Which is the mug?

Which is the cup?

Put the cup on the table on the left of the jug.

What do you use these objects for?

Without touching the objects show me how you would use them.

Mime how you would use these objects.

Touching the objects, show me how you would use these objects.

Touching the objects, show me how you would use them.

What colour is the jug?

What colour is the mug?

What colour is the cup?

Pour some drink from the jug into the cup and drink from the cup.

Drink from the cup.

Can you taste the drink?

Describe how it tastes.

Put the cup on the table on the right of the jug.

Is that your right?

Put the cup in front of the jug.

Put the cup behind the jug.

Task 4: Putting on a Shirt

I am going to ask you some questions and ask you to do something familiar. I want to see what you can do and what problems you may have, so that we can plan your treatment. Don't worry if you feel you cannot manage everything.

Please close your eyes. I am putting an object in your hand, and I want you to tell me what it is without looking.

What can you see on the table?

Which is the shirt?

Which is the button?

Put the button on the right of the shirt,

Now, put the button on the left of the shirt.

What do you use these objects for?

Without touching the objects show me how you would use them.

Mime how you would use these objects.

What colour is the button?

What colour is the shirt?

What colour is the jacket?

Put on this shirt.

Put on this blouse.

Put on this jacket.

Put on this cardigan.

Glossary

Abnormal tone: Including flaccidity, spasticity, tremor, rigidity, decreased strength and ataxia.

Agnosia: Inability to recognize familiar objects in spite of intact sensory capabilities (Crepeau, Cohn and Boyt Schell, 2009, p. 1154).

Anosognosia: Inability to recognise a part of one's own body (Grieve and Gnanasekaran, 2008, p. 224).

Aphasia (dysphasia): The inability or difficulty to express oneself through speech (expressive aphasia) or to comprehend the written or spoken word (receptive aphasia).

Apraxia: Inability to perform motor activities although sensory motor function is intact and the individual understands the requirements of the task (Crepeau, Cohn and Boyt Schell, 2009, p. 1154).

Ataxia: Loss of coordination and smoother interplay between muscles in the cerebellum due to damage leading to uncontrolled jerky movement (Edmans, 2010, p. 229).

Attention: The cognitive ability to focus on a task, issue or object (Crepeau, Cohn and Boyt Schell, 2009, p. 1154).

Attention span: The length of time an individual can be engaged in a specific activity

Auditory agnosia: Inability to recognize familiar sounds although hearing is intact.

Bilateral integration: The ability to perform purposeful movement that requires interaction between both sides of the body in a smooth and refined manner (Christiansen and Baum, 1991).

Body image: Subjective perception of the appearance of one's own body (Grieve and Gnanasekaran, 2008, p. 225).

Body scheme: Perception of the relative position of the body parts (Grieve and Gnanasekaran, 2008, p. 225).

Cognition: Mental processes including attention, memory, motivation, emotional control, motor control, sensory processing and thinking; the ability to think and to reason to solve problems (Crepeau, Cohn and Boyt Schell, 2009, p. 1155).

Colour agnosia: Inability to recognise familiar colours although vision is intact.

Comprehension: Mental grasp of meaning and relationships in language - the understanding of familiar spoken and written words, gestures and diagrams.

Constructional apraxia: Difficulty in the organisation of complex actions in two- or three-dimensional space (Grieve and Gnanasekaran, 2008, p. 226).

Depth and distance perception: The ability to use various visual cues to determine distance of objects from each other and self and to determine changes of planes of surfaces.

Dressing apraxia: Inability to dress oneself, primarily due to a disorder of spatial perception and/or body scheme (Grieve and Gnanasekaran, 2008, p. 226).

Expression: The ability to communicate through spoken and written word and gesture.

Figure-ground: The isolation of a shape or an object from its background (Grieve and Gnanasekaran, 2008, p. 227).

Fine motor coordination/dexterity: Smooth and harmonious action of groups of muscles working together to produce a desired finely controlled motion (Crepeau, Cohn and Boyt Schell, 2009, p. 1158).

Form constancy: The ability to attend to subtle variations in form (Zoltan et al, 1986).

Hearing acuity: The ability to hear clearly and distinctly, i.e. the capacity to hear pure tones at different frequencies and to discriminate speech sounds.

Ideational apraxia: A breakdown in the ability to perform a task because of a loss of neuronal model or mental representation of the procedure required for performance (Gillen, Glen and Burkhardt, Ann, 2004, p. 693).

Ideomotor apraxia: The inability to perform a task on command and to imitate gestures, even though the patient understands the concept of the task and is able to carry out habitual tasks automatically (O'Sullivan and Schmitz, 2007, p. 1337).

Initiation: The ability to start, either through physical or mental engagement, a task (Zoltan et al, 1986; Christiansen and Baum, 1991).

Kinesthesia: Sensation and awareness of active or passive movement (O'Sullivan and Schmitz, 2007, p. 1338).

Long-term memory: Memory that stores and processes information over periods of time from a few minutes to many years (Grieve and Gnanasekaran, 2008, p. 228).

Memory: Ability to register, retain, and recall past experience, knowledge, and sensation (Crepeau, Cohn and Boyt Schell, 2009, p. 1161).

Metamorphosia: is a perceptual distortion of objects in which the person reports that linear objects appear curved or discontinuous (American Academy of Ophthalmology, 2019).

Motor apraxia: Loss of access to kinaesthetic memory patterns that leads to an inability to perform purposeful movement because of defective planning and sequencing of movements, even though the idea and the purpose of the task is understood (Gillen, Glen and Burkhardt, Ann, 2004, p. 694).

Orientation: Awareness of self in relation to time, place, and identification of others (Crepeau, Cohn and Boyt Schell, 2009, p. 1163).

Perception: Integration of impressions from the different sensory sources into psychological, meaningful information. Meaning that the brain gives to sensory input; consists of 'maps' of every part of our body that are stores in the nervous system as the sensations from the skin, muscles,

joints, and gravity and movement receptors are organized and sorted during the person's daily activities (Crepeau, Cohn and Boyt Schell, 2009, p. 1164).

Perseveration: Unnecessary and prolonged repetition of a word, phrase or movement (Crepeau, Cohn and Boyt Schell, 2009, p. 1164).

Position in space: The ability to understand, interpret and deal with concepts of spatial position of objects in relation to each other and to self, such as up/down, in/out, in front/behind.

Position sense: The awareness of the body's position.

Proprioception: Sensations derived from movement (i.e., speed, rate, sequencing, timing, and force) and joint position; derived from stimulation to muscle and, to a lesser extent, joint receptors, especially from resistance to movement (Crepeau, Cohn and Boyt Schell, 2009, p. 1164).

Right / left discrimination: The ability to discriminate between left and right with relation to the sides of the body or body parts and to the external environment.

Sensation: A feeling or awareness that results from stimulation of the body's sensory receptors and transmission of the nerve impulse along an afferent fibre to the brain (O'Sullivan and Schmitz, 2007, p. 1345).

Short-term memory: Retention span of events, objects or ideas in immediate awareness (Zoltan et al, 1986); the ability to remember current day-to-day events, learn new information, and retrieve information after an interval of minutes, hours or days (O'Sullivan and Schmitz, 2007, p. 1345).

Simultanognosia: is a disorder of visual attention that involves impairment in interpreting a visual stimulus as a whole; scenes and objects are perceived in a piecemeal manner (Dalrymple et al, 2013, p. 1).

Somatognosia: Inability to perceive how the body parts relate to each other, and their relative positions in space (Grieve and Gnanasekaran, 2008, p. 231).

Space perception: The perception of a three-dimensional world of objects with perceived form and localization.

Spatial relations disorder: A constellation of deficits that have in common a difficulty in perceiving the relationship between objects in space, or the relationship between the self and two or more objects; included are disorders of figure-ground discrimination, form discrimination, spatial relations, position in space perception and topographic orientation (O'Sullivan and Schmitz, 2007, p. 1346).

Stereognosis: Recognition of familiar objects and forms and their shapes and sizes through touch.

Tactile agnosia (astereognosis): Inability to recognise familiar objects by touch although sensation is intact.

Tactile discrimination: The reception and comprehension of stimuli including two-point discrimination, deep pressure, vibration, localization of light touch stimuli and temperature changes.

Taste discrimination: The reception and comprehension of flavours.

Temperature discrimination: The reception and comprehension of warm and cold stimuli.

Unilateral body neglect (inattention): Failure to report, respond or orient to a unilateral stimulus presented to the body side contralateral to a cerebral lesion.

Visual acuity: Extent of visual perception dependent on the clarity of retinal focus, integrity of nervous elements and cerebral interpretation of the stimulus, i.e. the ability to perceive depth, shape and detail and to discriminate between light and dark. May be assessed using a Snellen chart or can refer to the smallest perceptible width of black lines on a white background.

Visual agnosia: Inability to recognise familiar objects and forms although vision is intact.

Visual attention: Voluntary act of visual fixation, focused gaze (Zoltan et al, 1986).

Visual field: The area within which objects can be seen. Normal visual field is approximately 60° upward, 60° inward, 70-75° downward and 100-110° outward.

Visual field loss: Loss to all or part of the visual Held including: homonomous hemianopia - blindness of one entire visual field, bitemporal hemianopia - loss of vision of both temporal fields, quadrantic anopia - loss of only a portion of the visual field (Kolb and Whishaw, 1990).

Visual neglect: Neglect of a portion of the visual field in addition to or in the absence of visual field deficit (Zoltan et al, 1986).

Visual scanning: Taking in the environment through ocular movements.

Visual spatial agnosia: A deficit in perceiving spatial relationships between objects and self (Zoltan et al, 1986).

Further Reading and Other Assessments

(See References for full details)

Area	Further reading Author, date, test reviews.	Other assessments Title, author, date.
Abnormal tone	Farber (1991) pp. 274 and 513.	*Motricity Index and Trunk Control Test* (Wade *et al*, 1985).
Agnosia	Zoltan *et al* (1986) pp. 91-4.	
Anosognosia	Zoltan *et al* (1986) pp. 64-5.	
Apraxia	Zoltan *et al* (1986) pp. 37-50. Edwards *et al* (1991).	*Cambridge Apraxia Battery* (Fraser and Turton, 1986).
Attention	Zoltan *et al* (1986) pp. 110-3. Duchek (1991a) pp.286-8, (1991b) p.529.	
Auditory agnosia	Zoltan *et al* (1986) pp. 95-6. Dunn (1991) pp. 491-4.	Free field hearing test (Macphee *et al,* 1988).
Colour agnosia	Zoltan *et al* (1986) p.92. RPAB review see Laver (1990).	Colour recognition sub-test of the *Rivermead Perceptual Assessment Battery* (RPAB) (Whiting *et al,* 1985).
Colour vision	Spence, A. (1989).	Colour blindness: Dvorine Pseudo-Isochromatic Plates (Dvorine, 1944).
Constructional apraxia	Zoltan *et al* (1986) pp. 37-44. Edwards *et al* (1991). COTNAB. For 65yr + standards see Laver and Huchison (1993).	Constructional ability sub-tests on the *Chessington OT Neurological Assessment Battery* (COTNAB) (Tyerman *et al*, 1986).
Depth/distance perception	Zoltan *et al* (1986) pp. 86-7.	
Figure-ground discrimination	Zoltan *et al* (1986) pp. 73-6. Review of RPAB: Laver (1990). COTNAB. For 65yr + standards see Laver and Huchison (1993).	Figure-ground sub-test of the *Rivermead Perceptual Assessment Battery* (RPBA) (Whiting *et al*, 1985). Overlapping and hidden figures subtests on COTNAB (Tyerman *et al*, 1986).
Form constancy	Zoltan *et al* (1986) pp. 76-8.	
Ideational and ideomotor apraxia	Zoltan *et al* (1986) pp. 46-50. Edwards *et al*, (1991).	*Cambridge Apraxia Battery* (Fraser and Turton, 1986).
Initiation		*Kitchen Task Assessment* (KTA) (Baum and Edwards, 1993). *Assessment of Motor Process Skills* (AMPS) (Fisher, 1992).
Language	Duchek (1991) pp. 294-7, 530. Gravell (1990) pp.41-53. Stevens (1990) pp.104-6. COTNAB. For 65yr + standards see Laver and Huchison (1993).	*Aphasia Screening Test* (Whurr, 1974). Ability to follow instructions sub-tests of the COTNAB (Tyerman *et al*, 1986).
Memory	Zoltan *et al* (1986) pp. 113-21.	*Behavioural Memory Test* (BMT)

	Duchek, (1991a) pp. 288-94, (1991b) pp. 529-30. Review of BMT and Fuld see Nelson (1990a and 1990b). Review of Kendrick see Bender (1990b).	(Wilson *et al*, 1985). *Fuld Object Memory Test* (Fuld, 1974). *Kendrick Cognitive Tests for the Elderly* (Kendrick, 1972).
Motor apraxia	Zoltan *et al* (1986) pp. 46-50.	
Orientation	Review of CAPE see Pattie (1988) and Bender (1990a).	*The Clifton Assessment Procedures for the Elderly* (CAPE) *– Cognitive Assessment Scale* (CAS) (Patties and Gilleard, 1979). *Middlesex Elderly Assessment of Mental State* (MEAMS) – first sub-test (Golding, 1989).
Perseveration	Zoltan *et al* (1986).	
Position in space	Zoltan *et al* (1986) pp. 78-80.	
Proprioception	Dunn (1991) pp. 482-5.	
Right/left discrimination	Zoltan *et al* (1986) pp. 65-7.	
Spatial Relations	Zoltan *et al* (1986) pp. 81-4.	
Somatognosia	Zoltan *et al* (1986) pp. 53-8.	
Tactile agnosia	Zoltan *et al* (1986) pp. 93-4.	
Tactile discrimination	Dunn (1991), pp. 474-82. COTNAB. For 65yr + standards see Laver and Huchison (1993).	Stereognosis and tactile discrimination sub-test of the *Chessington OT Neurological Assessment Battery* (COTNAB) (Tyerman *et al,* 1986).
Visual attention	Zoltan *et al* (1986) pp. 21-2.	
Visual deficits including visual field loss	Zoltan *et al* (1986) pp. 21-33. Dunn (1990) pp. 488-91.	
Visual neglect and unilateral neglect	Zoltan *et al* (1986) pp. 31-3. Zoltan *et al* (1986) pp. 59-64.	*Rivermead Behavioural Inattention Test* (BIT) (Wilson *et al*, 1987).
Visual object agnosia	Zoltan *et al* (1986) pp. 91-2.	
Visual perception	Zoltan *et al* (1986). RPAB review: Laver (1990). COTNAB. For 65yr + standards see Laver and Huchison (1993).	RPAB (Whiting *et al*, 1985). Visual perception sub-tests of COTNAB (Tyerman *et al*, 1986) and MEAMS (Golding, 1989).
Visual scanning	Zoltan *et al* (1986) pp. 22-5.	Sub-tests of the *Rivermead Behavioural Inattention Test* (BIT) (Wilson *et al*, 1987).
Visual spatial agnosia	Zoltan *et al* (1986) p. 93.	

References

AMERICAN ACADEMY OF OPHTHALMOLOGY (2019) *Metamorphopsia.* [On-line] Available from: https://www.aao.org/BcscSnippetDetail.aspx?id=577e2a33-77b5-43e4-9691-a817cf06a0b9 [accessed 4.10.2019]

ANASTASI, A. (1988). *Psychological Testing – Sixth Edition*. London: Collier Macmillan Publishers.

ARBESMAN, M., LIEBERMAN, D., and METZLER, C. A. (2014). Using evidence to promote the distinct value of occupational therapy. *American Journal of Occupational Therapy, 68* (4), 381-386. doi:10.5014/ajot.2014.684002

AVEYARD, H., and SHARP, P. (2013). *A beginner's guide to evidence-based practice in health and social care* (2nd ed). Maidenhead: McGraw-Hill / Open University Press.

BARTRAM, D. (1990). 'Reliability and validity'. In: BEECH, J.R. and HARDING, L. (Eds) (1990). *Assessment of the Elderly*. Windsor: NFER-NELSON.

BAUM, C.M. and EDWARDS, D.F. (1993). 'Cognitive performance in senile dementia of the Alzheimer's type: the kitchen task assessment'. *American Journal of Occupational Therapy*, 47, 431-6.

BAUM, C.M. and WOLF, T. M. (2013). *Executive Function Performance Test (EFPT) Manual*. Available from: http://www.ot.wustl.edu/about/resources/executive-function-performance-test-efpt-308 [accessed 15 August 2016].

BENDER, M.P. (1990a). 'Test review – Clifton Assessment Procedures for the Elderly'. In: BEECH, J.R. and HARDING, L. (Eds) (1990). *Assessment of the Elderly.* Windsor: NFER-NELSON.

BENNETT, S., HOFFMANN, T., McCLUSKEY, A., McKENNA, K., STRONG, J., and TOOTH, L. (2003). Introducing OTseeker (Occupational Therapy Systematic Evaluation of Evidence): a new evidence database for occupational therapists. *American Journal of Occupational Therapy, 57* (6), 635-638

BENNETT, A. E. and RITCHIE, K. (1975). Questionnaires in Medicine - A guide to their design and use. London: Oxford University Press.

BHAVNANI, G., COCKBURN, J., WHITING, S. and LINCOLN, N. (1983). 'The reliability of the Rivermead Perceptual Assessment Battery (RPAB)'. *British Journal of Occupational Therapists*, 46, 17-9.

BONDER, B. (1990) The Foundation - Standardized Assessments: Ethical Principles for Use. In: *Readings in Research.* Rockville: American Occupational Therapy Foundation, Inc.

CHRISTIANSEN, C. (1991). Occupational Therapy - Intervention for Life Performance. In: C. Christiansen, & C. Baum. (Eds). (1991). *Occupational Therapy - Overcoming Human Performance Deficits. New* Jersey: Slack.

CHRISTIANSEN, C. and BAUM, C. (Eds) (1991). *Occupational Therapy – Overcoming Human Performance Deficits.* New Jersey: Slack.

CICERONE, K. D., and TUPPER, D. E. (1986). Cognitive assessment in the neuropsychological rehabilitation of head-injured adults. In B. P. Uzzell & Y. Gross (Eds.), *The Clinical Neuropsychology of Intervention* (pp. 59-84). Boston: Martinus Nijhoff.

CLARKE, C., SEALEY-LAPES, C., and KOTSCH, I. (2001). *Outcome measures: information pack for occupational therapy*. London: College of Occupational Therapists.

COLLEGE of OCCUPATIONAL THERAPISTS (COT; 2003). *Occupational Therapy defined as a complex intervention.* London: COT.

COLLEGE of OCCUPATIONAL THERAPISTS (COT; 2004). *Guidance on the use of the International Classification of Functioning, Disability and Health (ICF) and the Ottawa Charter for Health Promotion in occupational therapy services.* London: COT.

COTRUS, A., and STANCIU, C. (2014). A Study on Dynamic Assessment Techniques, as a Method of Obtaining a High Level of Learning Potential, Untapped by Conventional Assessment**.** *Procedia - Social and Behavioral Sciences,* 116, 2616-2619. doi:10.1016/j.sbspro.2014.01.622

CRAINE, J. (1982). 'Principles of cognitive rehabilitation', In: TREXLER, L.E. (Ed). *Cognitive Rehabilitation: Conceptualization and Intervention.* New York: Plenum Press.

CREPEAU, E.B., COHN, E. S. and BOYT SCHELL, B.A. (2009) *Willard and Spackman's Occupational Therapy (11th ed).* Philadelphia: Lippincott, Williams & Wilkins.

CROCKER, L. and ALGINA, J. (1986). *Introduction to Classical and Modern Test Theory*. London: Holt, Rinehart and Winston, Inc.

DALRYMPLE, K.A., BARTON, J.J.S., and KINGSTONE, A.5(2013) A world unglued: simultanagnosia as a spatial restriction of attention. *Frontiers in Human Neuroscience*, 7 (145), 1-10. doi: 10.3389/fnhum.2013.00145

DOUCET, B. M. (2012). Neurorehabilitation: are we doing all that we can? *American Journal of Occupational Therapy,* 66 (4), 488-493. doi:10.5014/ajot.2012.002790

DOUGLAS, A., LETTS, L., and LIU, L. (2008). Review of Cognitive Assessments for Older Adults. *Physical & Occupational Therapy in Geriatrics,* 26 (4), 13-43. doi: 10.1080/02703180801963758

DUCHEK, J. (1991a). 'Cognitive dimensions of performance'. In CHRISTIANSEN, C. and BAUM, C. (Eds) (1991). *Occupational Therapy – Overcoming Human Performance Deficits.* New Jersey: SLACK.

DUCHEK, J. (1991b). 'Assessing Cognition' In CHRISTIANSEN, C. and BAUM, C. (Eds) (1991). *Occupational Therapy – Overcoming Human Performance Deficits.* New Jersey: SLACK.

DUNN, W. (1991). 'Assessing sensory performance enablers'. In CHRISTIANSEN, C. and BAUM, C. (Eds) (1991). *Occupational Therapy – Overcoming Human Performance Deficits.* New Jersey: SLACK.

DVORINE, I. (1944) *Dvorine Pseudo-Isochromatic Plates, (2nd ed)*. USA: The Psychological Corporation.

EBSCO HEALTH (2015). *CINAHL Database*. Available from: https://health.ebsco.com/products/the-cinahl-database. [Accessed 13 November 2016]

EBSCO HOST (2015). *A Definitive Guide to Complementary Medicine and Alternative Treatments*. Available from: https://www.ebscohost.com/academic/AMED-The-Allied-and-Complementary-Medicine-Database [Accessed 13 November 2016]

EDMANS, J. (2010) *Occupational Therapy and Stroke* (2nd ed). Oxford: Blackwell Publishing.

EDWARDS, D.F., DEUEL, R.K., BAUM, C.M. and MORRIS, J.C. (1991). 'A quantitative analysis of apraxia in senile dementia of the Alzheimer type: stage-related differences in prevalence and type', *Dementia*, 2, 142-9.

FALIK, L. H., & FEUERSTEIN, R. (2005). The current state and development of the Learning Propensity Assessment Device (LPAD). *Transylvanian Journal of Psychology: Special Issue on Dynamic Assessment,* 1, 17-56. Available from: http://epsz.pszichologia.psiedu.ubbcluj.ro/archive/TJP13.pdf [Accessed 13 November 2016]

FARBER, S.D. (1991). 'Assessing neuromotor performance enablers', In CHRISTIANSEN, C. and BAUM, C. (Eds) (1991). *Occupational Therapy – Overcoming Human Performance Deficits*. New Jersey: SLACK.

FEINSTEIN, A. R., JOSEPHY, B. R. and WELLS, C. K. (1986). Scientific and Clinical Problems in Indexes of Functional Disability. *Annals of Internal Medicine, 105*, 413 - 420.

FEUERSTEIN, R., FALIK, L. H., & FEUERSTEIN, R. (1995). *Revised LPAD examiner's manual*. Jerusalem: ICELP.

FEUERSTEIN, R., FALIK, L. H., & FEUERSTEIN, R. (2003). The Learning Potential Assessment Device In*: Feuerstein's Theory and Applied Systems: A Reader* (pp.51-97). Jerusalem: The International Centre for the Enhancement of Learning Potential (ICELP).

FEUERSTEIN, R., FALIK, L. H., & FEUERSTEIN, R. (2002). *The Dynamic Assessment of Cognitive Modifiability: The LPAD***.** Jerusalem: ICELP Press.

FEUERSTEIN INSTITUTE (n.d.) *Learning Propensity Assessment Device (LPAD)*. Available from: http://acd.icelp.info/what-we-teach/learning-propensity-assessment-device.aspx [Accessed 13 November 2016]

FISHER. A.G. (1992). 'The assessment of motor process skills – research edition 6.1'. Unpublished – provided by author.

FISHER, A. G. and SHORT-DeGRAFF, M. (1993). Improving Functional Assessment in Occupational Therapy: Recommendations and Philosophy for Change. *American Journal of Occupational Therapy*, *47,* 199 - 201.

FRASER, C.M. and TURTON, A. (1986). 'The development of the Cambridge Apraxia Battery'. *Occupational Therapy*, August, 248-52.

FRY, K., and PERLMUTTER, M.S. (1991). 'The IWJ functional assessment'. Washington University School of Medicine, St Louis: unpublished – provided by the authors.

FULD, P.A. (1974). *Fuld Object Memory Test*. Windsor: NFER- NELSON.

GILLEN, G., and BURKHARDT, A. (2004) *Stroke Rehabilitation a Function-Based Approach* (2nd ed). Missouri: Mosby.

GOLDING, E. (1989). *The Middlesex Elderly Assessment of Mental State*. Suffolk: Thames Valley Test Company.

GOLISZ, K. M., and TOGLIA, J. P. (2003). Perception and cognition. In: Crepeau, E. B., Cohn, E. S., & Boyt Schell, B. A (Eds). *Willard and Spackman's occupational therapy (pp.395-411)* (10th ed.) Philadelphia: Lippincott Williams and Wilkins.

GOLLEDGE, J. (2006) Enabling participation in occupations after stroke. In: Addy, L. M. (Ed.), *Occupational Therapy Evidence in Practice for Physical Rehabilitation* (pp.199-230). Oxford: John Wiley and Sons

GRAVELL, R. (1990). 'Assessment of communication'. In: BEECH, J.R. and

GURALMIK, J.M., BRANCH, L.G., CUMMINGS, S.R. and CURB, J.D. (1989). 'Physical performance measures in aging research'. *Journal of Gerontology*, 44, 151-6.

GRIEVE, J., and GNANASEKARAN, L. (2008) *Neuropsychology of Occupational Therapists* (3rd ed). Oxford: Blackwell Publishing.

HADAS-LIDOR, N. and WEISS, P. (2005). Dynamic cognitive intervention: Application in occupational therapy. In N. Katz (Ed.) *Cognition and occupation across the life span: Models for intervention in occupational therapy* (pp.323-350). Bethesda, MD: AOTA Press.

HADAS-LIDOR, N., WEISS, P., and KOZULIN, A. (2011). Dynamic cognitive intervention: Application in occupational therapy. In N. Katz (Ed.) *Cognition and occupation across the life span: neuroscience, neurorehabilitation and models of intervention in occupational therapy* (pp.323-350) (3rd ed). Bethesda, MD: AOTA Press.

HARDING, L. (Eds) (1990). *Assessment of the Elderly*. Windsor: NFER-NELSON.

HAYNES, S. N., RICHARD, D. C. S., and KUBANY, E. S. (1995). Content Validity in Psychological Assessment: A Functional Approach to Concepts and Methods. *Psychological Assessment*, 7 (3), 238-247. Doi: 10.1037/1040-3590.7.3.238.

HAYWOOD, H. C., and LIDZ, C. S. (2007). *Dynamic Assessment in Practice: clinical and educational applications.* Cambridge: Cambridge University Press.

HAYWOOD, H. C., and MILLER, M. B. (2003). Dynamic assessment of adults with traumatic brain injuries. *Journal of Cognitive Education and Psychology Online* 3, 137-158. doi: 10.1891/194589503787383046

HESSELS-SCHLATTER, C. (2002). A dynamic test to assess learning capacity in people with severe impairments. *American Journal on Mental Retardation*, 107, 340-351

HERSCH, G. I., LAMPORT, N. K., and COFFEY, M. S. (2005). *Activity analysis: application to occupation* (5th ed). Thorofare, SLACK Incorporated.

KATZ, N., LIVNI, L., BAR-HAIM EREZ, A., and AVERBUCH, S. (2011). *Dynamic Lowenstein Occupational Therapy Cognitive Assessment (DLOTCA)*. Pequannock, NJ: Maddak.

KATZ, N., BAR-HAIM EREZ, A., LIVNI, L., and AVERBUCH, S. (2012). Dynamic Lowenstein Occupational Therapy Cognitive Assessment**:** evaluation of potential to change in cognitive performance. *American Journal of Occupational Therapy,* 66 (2), 207-215. doi:10.5014/ajot.2012.002469

KATZ, N., AVERBUCH, S., and BAR-HAIM EREZ, A. (2012). Dynamic Lowenstein occupational therapy cognitive assessment-geriatric version (DLOTCA-G): assessing change in cognitive performance. *AJOT: American Journal of Occupational Therapy*, 66 (3), 311- 319. doi:10.5014/ajot.2012.002485.

KENDRICK, D. (1990). Assessing memory. In In: BEECH, J.R. and HARDING, L. (Eds) (1990). *Assessment of the Elderly*. Windsor: NFER-NELSON.

KENDRICK. D. (1972). *Kendrick Cognitive Tests for the Elderly*. Windsor: NFER-NELSON.

KLINE, P. (1986). *A Handbook of Test Construction*. London: Methuen

KOENIGSKNECHT, C., and SMITH, H. (2012). *Lowenstein Occupational Therapy Cognitive Assessment for Adults*. Availablefrom: http://www.otshow.com/?filename=lowenstein-occupational-therapy-cognitive-assessment-for-adults. [12 November 2016]

KOLB, B. and WHISHAW, I.Q. (1990). *Fundamentals of Human Neuropsychology (3rd ed)*. San Francisco: W.H. Freeman and Company.

LAVER, A.J. (1994). The development of the Structured Observational Test of Function. Ph.D thesis. Guildford: Department of Psychology, University of Surrey. (Available from the College of Occupational Therapists library in London, UK)

LAVER, A.J. (1990). 'Test review – Rivermead Perceptual Assessment Battery'. In BEECH, J.R. and HARDING, L. (Eds) (1990), *Testing people*. Windsor: NFER-NELSON.

LAVER, A.J. and HUCHISON, S. (1993). *The Performance and Experience of Normal Elderly People on the Chessington Occupational Therapy Neurological Assessment Battery (COTNAB) – 65yr+ Standards*. Nottingham: Nottingham Rehab.

LAVER, A.J. and BAUM, C.M. (1992). An assessment model. Unpublished, St Louis: Program in Occupational Therapy, Washington University School of Medicine.

LAVER, A. J. and POWELL, G. E. (1995). *The Structured Observational Test of Function (SOTOF).* Windsor: NFER-NELSON.

LAW, M. (1993). 'Evaluating activities of daily living: directions for the future'. *American Journal of Occupational Therapy*, 47, 233-7.

LAW, M. (1997). Self Care. In: Van Deusen, J., & Brunt, D. (Eds.) *Assessment in Occupational Therapy and Physical Therapy* (pp. 431). London: W B Saunders Company.

LAW, M. and LETTS, L. (1989). 'A critical review of scales of activities of daily living'. *American Journal of Occupational Therapy*, 43, 233-7.

LETTS, L., and BOSCH, J. (2001). Measuring Occupational Performance in Basic Activities of Daily Living. In: Law, M., Baum, C., & Dunn, W. (Eds.) *Measuring Occupational Performance: Supporting Best Practice in Occupational Therapy* (pp.121-159). Thorofare: Slack.

LIDDLE, J., and EAGLES, R. (2014). Moderate evidence exists for occupational therapy-related interventions for people with Parkinson's disease in physical activity training, environmental cues and individualised programmes promoting personal control and quality of life. *Australian Occupational Therapy Journal*, 61 (4), 287-288. doi: 10.1111/1440-1630.12153

LINCOLN, N.B. and CLARKE, D. (1987) 'The performance of normal elderly people on the Rivermead Perceptual Assessment Battery (RPAB)'. *British Journal of Occupational Therapy*, 50, 156-7.

LINCOLN, N.B. and EDMANDS, J.A. (1989). 'A shortened version of the Rivermead Perceptual Assessment Battery'. *Clinical Rehabilitation*, 3, 215-21.

McARTHUR, M. A., and SPALDING, N. J. (1997). Book reviews: the Structured Observational Test of Function. *Disability and Rehabilitation*, 19 (11), 501. doi: 10.3109/09638289709166849.

MACPHEE, G., CROWTHER, J. and McALPINE, C. (1988). 'A simple screening test for hearing impairment in elderly patients'. *Age and Ageing*, 17, 347-51.

MILLER, G.A. (1968). 'What is information measurement?' In: BUCKLEY, W. (1968). *Modern System Research for the Behavioural Scientist – A Sourcebook*. Chicago: Aldine Publishing Company, p.123.

MISSIUNA, C. (1987). Dynamic assessment: A model for broadening assessment in occupational therapy. *Canadian Journal of Occupational Therapy,* 54, 17-21.

MURDOCK, C. (1992). A Critical Evaluation of the Bartel Index, Part 1. *British Journal of Occupational Therapy, 55,*109 - 111.

NATIONAL CENTER FOR MEDICAL REHABILITATION RESEARCH (NCMRR) (1992). *Report and Plan for Medical Rehabilitation Research to Congress*. National Center for Medical Rehabilitation Research, Bethesda, Maryland: National Institute for Child Health and Human Development, National Institutes of Health.

NELSON, H. (1982). *National Adult Reading Test (NART).* Windsor: NFER-NELSON.

NELSON, H. (1990a). 'Test review of the Fuld Object Memory Evaluation', In: BEECH, J.R. and HARDING, L. (Eds) (1990). *Testing people*. Windsor: NFER-NELSON.

NELSON, H. (1990b). 'Test review of the Rivermead Behavioural Memory Test'. , In: BEECH, J.R. and HARDING, L. (Eds) (1990). *Testing people*. Windsor: NFER-NELSON.

NELSON, H. and WILLISON, J. (1991). *National Adult Reading Test (NART), (2nd ed)*. Windsor: NFER-NELSON.

NORUSIS, M. J. (1990). *SPSS Base System Users Guide.* Chicago: SPSS Inc.

NOTT, M. T., CHAPPARO, C., and HEARD, R. (2008). Effective occupational therapy intervention with adults demonstrating agitation during post-traumatic amnesia. *Brain Injury,* 22 (9), 669-683. doi: 10.1080/02699050802227170

OPACICH, K. J. (1991). Assessment and Informed Decision Making. In: C. Christiansen, & C. Baum. (Eds). (1991). *Occupational Therapy - Overcoming Human Performance Deficits.* New Jersey: Slack.

O'SULLIVAN, S., and SCHMITZ, T. (2007) *Physical Rehabilitation* (5th ed). Philadelphia: FA Davis Company.

OTTENBACHER, K.J. (1980). 'Cerebral vascular accident: some characteristics of occupational therapy forms'. *American Journal of Occupational Therapy*, 34, 268-71.

OTTENBACHER, K.J. and TOMCHEK, S.D. (1993). 'Reliability analysis in therapeutic research: practice and procedures'. *American Journal of Occupational Therapy*, 47, 10-6.

PATTIE, A. (1988). 'Measuring levels of disability – the Clifton Assessment Procedures for the Elderly'. In: WATTS, J.P. and HINDMARCH, I. (Eds) (1988). *Psychological Assessment of the Elderly*. London: Churchill Livingstone.

PATTIE, A. and GILLEARD, C. (1979). *Clifton Assessment Procedures for the Elderly.* London: Hodder and Stoughton.

PEDRETTI, L.W. (1985). *Occupational Therapy Practice Skills for Physical Dysfunction, 2nd ed.* St Louis: C V Mosby.

PORTNEY, L. G. and WATKINS, M. P. (1993). *Foundations of Clinical Research - Applications to Practice.* Connecticut: Appleton & Lange.

ROGERS, J.C. and HOLM, M.B. (1991). 'Occupational therapy diagnostic reasoning: a component of clinical reasoning'. *American Journal of Occupational Therapy*, 45, 1045-53.

ROGERS, J.C. and MASAGANTANI, G, (1982). 'Clinical reasoning of occupational therapists during the initial assessment of physically disabled patients'. *The Occupational Therapy Journal of Research*, 2, 195-219.

SANDERSON, H., and GITSHAM, N. (1991). *A holistic sensory approach: A guide to sensory stimulation for people who have profound learning disabilities* (2nd ed). London: David Fulton.

SLATER, R. (2013). *Use Google Scholar Effectively for Research*. Available from: http://www.library.illinois.edu/ugl/howdoi/use_google_scholar.html [accessed 13 November 2016]

SMITH, H. (2012). *Lowenstein Occupational Therapy Cognitive Assessment for Adults*. Retrieved from: http://www.otshow.com/?filename=lowenstein-occupational-therapy-cognitive-assessment-for-adults[accessed 13 November 2016]

SPENCE, A. (1989). *Biology of Human Ageing*. New Jersey: Prentice-Hall.

SPITZNAGEL, E. (1991). Statistics - Course Notes for Math 320 - Academic Year 1991 - 1992. St Louis: Department of Mathematics, Washington University.

STEVENS, S.J. (1990). 'Test review of the Aphasia Screening Test'. In: BEECH, J.R. and HARDING, L. (Eds) (1990). *Assessment of the Elderly*. Windsor: NFER-NELSON.

STERNBERG, R. J., and GRIGORENKO, E. L. (2002). *Dynamic testing: The nature and measurement of learning potential*. Cambridge: Cambridge University Press.

STEIN, F. (1989). *Anatomy of Clinical Research - An Introduction to Scientific Inquiry in Medicine, Rehabilitation, and Related Health Professions, Revised Edition.* New Jersey: Slack Inc.

TEXAS EDUCATION AGENCY (2011). *Hierarchy of cueing and prompting*. Austin, TX: Texas Education Agency.

TOGLIA, J. P. (1993). *Contextual Memory Test Manual*. USA: Pearson.

TOGLIA, J. P. (1994). *Dynamic assessment of categorization: The Toglia category assessment manual.* Pequannock, NJ: Maddak.

TOGLIA, J. P. (1998). A dynamic interactional approach to cognitive rehabilitation. In: Katz, N. (Ed.) *Cognition and occupation across the lifespan: models for intervention in occupational therapy* (pp.5-47). Bethesda: American Occupational Therapy Association.

TOGLIA, J. P. (2005). A dynamic interactional approach to cognitive rehabilitation. In: Katz, N. (Ed.) *Cognition and occupation across the lifespan: models for intervention in occupational therapy* (pp.29-72). Bethesda: American Occupational Therapy Association.

TOGLIA, J. P. (2011). The Dynamic Interactional Model of Cognition in Cognitive Rehabilitation. In: Katz, N. (Ed.) *Cognition and occupation across the life span: Models for intervention in occupational therapy* (pp.161-201) (3rd ed). Bethesda: AOTA Press.

TOGLIA, J., and CERMAK, S. (2009). Dynamic Assessment and Prediction of Learning Potential in Clients with Unilateral Neglect. *American Journal of occupational therapy,* 63 (5), 569 – 580. doi:10.5014/ajot.63.5.569

TOGLIA, J. P., GOLISZ, K., and GOVEROVER, Y. (2008). Evaluation and intervention for cognitive perceptual impairments. In: Crepeau, E. B., Schell, B., & Cohn, E. (Eds.) *Willard and Sparkman's Occupational Therapy* (pp.739-776) (11th ed). Philadelphia: Lippincott Williams and Wilkins.

TROMBLY, C. (1993). Anticipating the Future: Assessment of Occupational Function. *American Journal of Occupational Therapy*, *47,* 253 - 257.

TYERMAN, R., TYERMAN, A., HOWARD, P. and HADFIELD, C. (1986). *The Chessington O.T. Neurological Assessment Battery Introductory Manual.* Nottingham: Nottingham Rehab Limited.

UPRICHARD, S., KUPSHIK, G., PINE, K., & FLETCHER, B. (2009). Dynamic assessment of learning ability improves outcome prediction following acquired brain injury. *Brain Injury,* 23 (4), 278-290. doi: 10.1080/02699050902788444and

VYGOTSKY, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

WADE, D.T. et al. (1985). Stroke – *A Critical Approach to Diagnosis, Treatment and Management*. London: Chapman and Hall Medical.

WHITING, S. and LINCOLN, N. (1980). 'An A.D.L assessment for stroke patients'. *Occupational Therapy*, February, 44-6,

WHITING, S., LINCOLN, N., BHAVNANI, G. and COCKBURN, J. (1985). *RPAB – Rivermead Perceptual Assessment Battery – Manual.* Windsor: NFER-NELSON.

WHURR, R. (1974). *Aphasia Screening Test*. London: Whurr Publishers.

WIEDL, K. H., SCHOTTKE, H., GARCIA, M., and DOLORES, C. (2001). Dynamic assessment of cognitive rehabilitation potential in schizophrenic persons and in elderly persons with and without dementia. *European Journal of Psychological Assessment,* 17(2), 112-119. doi: 10.1027//1015-5759.17.2.112

WILSON, B., COCKBURN, J. and BADDELEY, A. (1985). *The Rivermead Behavioural Memory Test.* Bury St Edmunds: Thames Valley Test Company.

WILSON, B., COCKBURN, J. and HALLIGAN, P. (1987). *Rivermead Inattention Test.* Bury St Edmunds: Thames Valley Test Company.

ZOLTAN, B., SIEV, E. and FREISHTAT, B. (1986). *Perceptual and Cognitive Dysfunction in the Adult Stroke Patient – A manual for Evaluation and Treatment, (2nd ed).* New Jersey: Slack