ACRM
AMERICAN CONGRESS OF REHABILITATION MEDICINE

Archives of Physical Medicine and Rehabilitation

journal homepage: www.archives-pmr.org

Archives of Physical Medicine and Rehabilitation 2018;99:1848-75

REVIEW ARTICLE (META-ANALYSIS)

Systematic Review of Measurement Property Evidence for 8 Financial Management Instruments in Populations With Acquired Cognitive Impairment

Lisa Engel, PhD,[a,b] Adora Chui, MScOT,[b] Dorcas E. Beaton, PhD,[a,c,d,e] Robin E. Green, PhD,[a,f,g] Deirdre R. Dawson, PhD[a,b,c,f]

From the [a]Rehabilitation Sciences Institute, University of Toronto; [b]Rotman Research Institute, Baycrest Health Sciences; [c]Department of Occupational Science and Occupational Therapy, University of Toronto; [d]Institute for Work & Health; [e]Li Ka Shing Knowledge Institute, St. Michael's Hospital; [f]Toronto Rehabilitation Institute, University Health Network; and [g]Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada.

Abstract

Objectives: To critically appraise the measurement property evidence (ie, psychometric) for 8 observation-based financial management assessment instruments.

Data sources: Seven databases were searched in May 2015.

Study selection: Two reviewers used an independent decision-agreement process to select studies of measurement property evidence relevant to populations with adulthood acquired cognitive impairment, appraise the quality of the evidence, and extract data. Twenty-one articles were selected.

Data extraction: This review used the COnsensus-based Standards for the selection of health Measurement Instruments review guidelines and 4-point tool to appraise evidence. After appraising the methodologic quality, the adequacy of results and volume of evidence per instrument were synthesized. Measurement property evidence with high risk of bias was excluded from the synthesis.

Data synthesis: The volume of measurement property evidence per instrument is low; most instruments had 1 to 3 included studies. Many included studies had poor methodologic quality per measurement property evidence area examined. Six of the 8 instruments reviewed had supporting construct validity/hypothesis-testing evidence of fair methodologic quality. There is a dearth of acceptable quality content validity, reliability, and responsiveness evidence for all 8 instruments.

Conclusions: Rehabilitation practitioners assess financial management functions in adults with acquired cognitive impairments. However, there is limited published evidence to support using any of the reviewed instruments. Practitioners should exercise caution when interpreting the results of these instruments. This review highlights the importance of appraising the quality of measurement property evidence before examining the adequacy of the results and synthesizing the evidence.

Archives of Physical Medicine and Rehabilitation 2018;99:1848-75

© 2018 by the American Congress of Rehabilitation Medicine

Financial capital is intertwined with all aspects of adult life, and the ability to manage personal finances, such as banking or paying for goods and services, is a central part of autonomy and independence.[1] Yet, many rehabilitation practitioners and researchers working with people who have adult-onset cognitive impairments are unsure of how to aptly assess financial management (FM). The purpose of this review is to identify and critically analyze the measurement property (MP) evidence of 8 selected observation-based instruments related to FM, leading to evidence-based recommendations on their use in adult populations with cognitive impairment. The included instruments are:

Presented in part as a podium presentation to the Fifth Annual Occupational Therapy Summit of Scholars, May 21–22, 2016, Pittsburgh, PA (Dr. Engel); and to the Canadian Association of Occupational Therapists Annual Conference, May 27–30, 2015, Winnipeg, Manitoba, Canada (Drs. Engel, Beaton D, Green, and Dawson).

Supported by grants awarded by the Canadian Institutes of Health Research, Ontario Neurotrauma Foundation, and the Canadian Partnership Stroke Recovery, Baycrest Health Sciences, and the University of Toronto.

Disclosure statement: Dr. Engel reports scholarships and fellowships from Canadian Institutes of Health Research, Ontario Neurotrauma Foundation, Baycrest Health Sciences, and University of Toronto, during the conduct of the study; the remaining authors have nothing to disclose.

0003-9993/18/$36 - see front matter © 2018 by the American Congress of Rehabilitation Medicine
https://doi.org/10.1016/j.apmr.2018.02.004

Cognitive Competency Test (CCT); Everyday Functioning Battery/Functional Impact Assessment (EFB/FIA); Financial Capacity Instrument (FCI); Financial Competency Assessment Inventory (FCAI); Independent Living Scales (ILS); Kohlman Evaluation of Living Skills (KELS); Measurement for Awareness of Financial Skills (MAFS); and Semi-structured Clinical Interview of Financial Capacity (SCIFC).

Appendix 1 provides a description and citation for these instruments.

FM limitations are of increased rates in diagnostic populations that experience cognitive impairments such as adults with acquired brain injury, dementia, multiple sclerosis, and mental health populations.[2] There is a paucity of FM limitation prevalence data for cognitive impairment populations; however, 3 studies found 23% to 33% of people with brain injury self-report FM activity limitations.[3–5] Despite the potentially detrimental consequences of FM errors on the quality of life of adults with cognitive impairments and their dependents, relatively little research addresses how best to assess the life activity of FM.[6] However, rehabilitation practitioners are currently involved in assessing FM.[7] FM assessment purposes are descriptive (ie, discriminative) and evaluative (ie, outcome), as it is important for practitioners to correctly identify and monitor adults with FM limitations.

The potential consequences of FM errors, coupled with unemployment, financial exploitation, low socioeconomic status, and decreased available financial resources, which are all of increased prevalence in adult populations with cognitive impairments, make the correct assessment of FM limitations vital.[8–11] These factors can mean that adults with cognitive impairments may have less financial capital overall to mitigate financial errors if they do happen. Individuals with FM limitations who retain control of their finances can make FM decisions and actions that are economically detrimental to themselves and their dependents. Conversely, the incorrect assessment of FM limitations can also be a threat to well-being and quality of life. When adults are deemed unable to manage their finances (regardless of whether the designation is correct), their autonomy and participation in everyday life events are restricted.[12,13]

In our previous systematic search and content analysis, we identified FM assessment instruments with comprehensive items in multiple FM domains or task areas.[14] That work provided a foundational understanding of currently available instrument options. To endorse instrument(s) for use, the MP evidence needs to be critically evaluated by including reliability, validity, and responsiveness evidence. This includes appraising the methodologic quality of the MP evidence (ie, risk of bias) before examining the MP evidence results (ie, adequacy of results). Further, considering how current clinical and research practice have used FM instruments, reviewed instruments require MP evidence for both descriptive and evaluative measurement purposes and for both group (eg, research, program evaluation) and individual assessment.

Other reviews have examined instruments related to FM.[13,15–17] Although these reviews provide a foundational synthesis of some available FM instruments and MP evidence, none used systematic review methods or completed a critical appraisal of the methodologic quality of the MP evidence. Further, the evidence included was not specific to adult populations with cognitive impairment.

To further our understanding of FM assessments, the aims of this study are to: identify available MP evidence, the quality of the evidence, and the diagnostic populations for whom the evidence is applicable; determine MP evidence gaps for these instruments; and make recommendations about the use of these instruments. The objective is to systematically review the MP evidence applicable to populations with adult-onset cognitive impairment for eight FM-related instruments.

List of abbreviations:

CCT	**Cognitive Competency Test**
COSMIN	**COnsensus-based Standards for the selection of health Measurement Instruments research group**
EFB/FIA	**Everyday Functioning Battery/Functional Impact Assessment**
FCAI	**Financial Competency Assessment Inventory**
FCI	**Financial Capacity Instrument**
FM	**financial management**
ILS	**Independent Living Scales**
KELS	**Kohlman Evaluation of Living Skills**
MAFS	**Measurement for Awareness of Financial Skills**
MP	**measurement property**
SCIFC	**Semi-structured Clinical Interview of Financial Capacity**

Methods

We used assessment systematic review methods to identify applicable peer-reviewed published studies based on the guidelines from the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN; http://www.cosmin.nl)[18,19] and Preferred Reporting Items for Systematic Reviews and Meta-Analyses.[20] Reviewers followed a search protocol (available from the authors) that was developed a priori with the assistance of a University librarian experienced in databases and systematic reviews. The protocol included search, screening/selection, extraction, data synthesis processes, and review team roles.

Data Sources

We electronically searched 7 databases in May 2015: AgeLine (hosted by EBSCO Information Services); ASSIA (hosted by ProQuest); CINAHL (hosted by EBSCO); Embase (hosted by Ovid), MEDLINE (hosted by Ovid), PsycINFO (hosted by Ovid), and Scopus (hosted by Elsevier). These databases cover literature across multiple areas of study (eg, rehabilitation, psychology, gerontology, social sciences) and geographic areas (eg, North America, Europe). When possible, we consulted instrument and study authors to inquire about the volume of evidence, used database alerts to identify newer articles, and examined the references of included studies and related reviews to identify other applicable studies.

Search Terms

Search terms included population, instrument names and acronyms, and MP terms (appendix 2). We based the MP terms and filters on Terwee and colleagues' PubMed precise search structure for MP evidence reviews.[21] When databases allowed, we limited the searches by peer-reviewed publication; English language

Table 1 Reasons for poor COSMIN methodologic quality ratings*

No.[†]	Reason	No. of studies[‡]
Design issues across COSMIN measurement property areas		
1	Inappropriate statistics	n=3
2	Insufficient sample size	n=7
3	Study methods not adequately described	n=2
4	Other general design issues: dual role of measure developer and assessor; reliability analysis of healthy controls only; validity data of same participants measured multiple times; unexplained increased volume of missing data	n=7
Reasons specific to internal consistency		
5	Uni-dimensionality not checked or referenced	n=6
6	Internal consistency not checked for each subscale	n=1
Reasons specific to reliability		
7	Less than 2 measurements available or number of measurements not described	n=1
8	Measurements not independent of each other (eg, scoring/analysis based on videotaped measure administrations) or unsure of independence of measurements	n=3
9	Time interval not stated, inappropriate time interval between measures, unclear if test conditions were similar for both measurements, and/or unsure if participants stable between measurements	n=2
Reasons specific to content validity		
10	Lacks assessment of whether all items refer to relevant aspects of the construct to be measured	n=1
11	Lacks assessment of whether all items relevant from perspective of the study population (eg, age, gender, diagnostic characteristics, cultural influences may not be accounted for)	n=2
12	Lacks assessment of whether all items together comprehensively reflect the construct to be measured	n=1
Reasons specific to hypothesis testing or responsiveness		
13	Lacks *a* priori hypothesis statement(s)	n=2
14	For convergent validity, inadequate description of the comparator instrument and/or its measurement evidence	n=5

* Reasons for poor quality ratings within the included articles; based on COSMIN 4-point scale, but not inclusive of all COSMIN measurement property items that can have a poor rating.[27]

† Reason number relates to superscript numbers within Table 2.

‡ Total number of studies in this column exceeds 21 as some measurement property evidence per study has more than one reason for poor quality rating.

(secondary to the language abilities of the reviewers); human subjects; and adults 18 years of age and older. We filtered by irrelevant phrases related to the instrument acronyms to increase search precision (eg, “central corneal thickness” for CCT). The search was not limited by year of publication, as older literature may be relevant to our objective.[18]

Screening and Selection

Two reviewers (A.C. and L.E.) independently screened the titles, abstracts, and full articles according to a priori inclusion and exclusion criteria. Both reviewers were master’s-level rehabilitation professionals with experience in systematic reviews and assessment. At each level of review (ie, title, abstract, full article), reviewers reached agreement before continuing independent selection at the next level.

Articles were included if they discussed instrument development or MP evidence for one of the target instruments. Articles were excluded if they did not include measurement evidence of the target instruments or a population with possible cognitive impairments. The latter exclusion criterion did not apply to content validity studies in which financial or rehabilitation experts were participants. Conference abstracts and non−peer-reviewed published MP evidence from assessment manuals were excluded. Only peer-reviewed sources were considered owing to the increased bias of results when using standardized instruments without peer-reviewed published evidence.[22,23]

Critical Appraisal and Extraction

The same 2 reviewers again used an independent review followed by a discussion agreement process to rate MP evidence methodologic quality and to extract data. A third author (D.B.), with expertise in measurement, was consulted for disagreements. Reviewers pilot tested the quality rating and extraction process on 4 articles. Minor wording and process changes were made resulting in the final data extraction process.

The extraction and appraisal process included 5 main steps. First, reviewers identified the specific MP evidence examined in each study.[24,25] Most studies were rated according to the MP area reported by its authors. For the studies reportedly examining criterion validity, we rated them using the COSMIN area of hypothesis testing (ie, construct validity), as there is no ‘gold-standard’ (i.e., totally valid instrument) FM comparator instrument or measure available, which is needed to assess criterion validity.[18]

Second, we separated the included articles into categories of direct and indirect MP evidence.[26] Direct evidence studies had a specific purpose to describe instrument development or to analyze the instrument’s MP evidence; indirect evidence studies did not have this specific purpose or used the target instrument as the comparator to validate a different instrument.[18] This review focused on analyzing and synthesizing the direct evidence studies for the 8 FM instruments.

Third, raters analyzed the methodologic quality of each MP evidence area using the standardized COSMIN 4-point scale,

rating each area of MP evidence discussed within each study as poor, fair, good, or excellent.[27] MP areas in the included studies rated as poor methodologic quality were excluded owing to the high risk of bias (see table 1 for poor score reasons).

Fourth, raters extracted and analyzed the adequacy of results of the MP evidence with fair, good, or excellent COSMIN methodologic ratings. We used the following guidelines to judge the adequacy of results[18,25,28]: internal consistency (Cronbach α): 0.70 to 0.95 acceptable; reliability (intraclass correlation coefficient [ICC]): 0.70 acceptable, >0.80 strong; hypothesis testing/convergent construct validity and responsiveness: correlations were compared to a priori hypotheses that included statements about expected correlation magnitudes. In the absence of a priori hypotheses or stated absolute magnitudes for correlations, we used preset statistical criteria to determine correlation strength for each hypothesis stated (ie, $r<0.25$ weak, $r=0.25\text{-}0.50$ fair, $r=0.50\text{-}0.75$ good/moderate, and $r>0.75$ excellent/strong).

Last, we calculated 2 outcome interpretability values when test-retest reliability and baseline standard deviation (SD) were available. We determined standard error of the measurement (SEM_{diff}) using the reported test-retest reliability ICC to inform measurement error. We then calculated the 95% confidence interval for the smallest detectable change (SDC; also known as the minimal detectable change) to indicate the interpretability of change (ie, change that can be attributed beyond measurement error).[18,29] We used the following equations[29,30]:

$$SEM_{diff} = SD_{baseline} * \sqrt{[2 * (1 - ICC_{test-retest})]},$$

$$SDC_{95} = 1.96 * SEM_{diff}.$$

Evidence synthesis

Because COSMIN guidelines do not provide direction for synthesizing MP evidence, we studied the overall volume of each MP evidence area across all included studies per instrument using the Cochrane Back Review Group guidelines[31]: strong evidence: results from one study of excellent methodologic quality or consistent results from multiple good studies; moderate evidence: results from one study of good methodologic quality or consistent results from multiple fair studies; limited evidence: results from one study of fair methodologic quality; conflicting evidence: results not consistent across studies of fair, good, or excellent quality; and unknown evidence: no published studies or evidence only from studies of poor methodologic quality.

Results

We found 404 titles (fig 1). Through the screening and selection process, 21 studies of direct measurement evidence and 21 studies of indirect evidence were found (see appendices 3 and 4 for study citation and results summaries). All 8 instruments had studies regarding construct validity hypothesis testing, 5 had studies

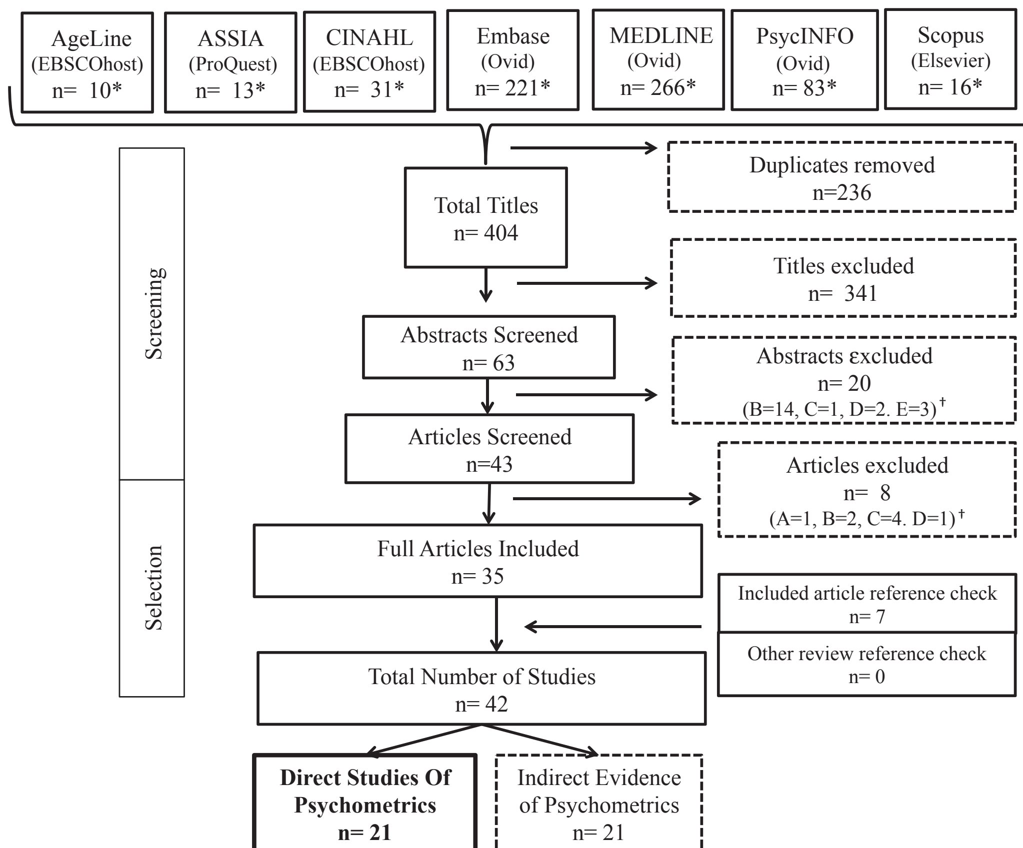

Fig 1 Flowchart of search, screening, and selection. *Duplicates removed per database. †Reasons for exclusion: A, article not full manuscript; B, target instrument not included; C, population not adult-onset cognitive impairment; D, study does not provide measurement property evidence; E, duplicate articles.

Table 2 COSMIN Methodological Quality Ratings of Measurement Properties for Each Instrument per Each Study of Direct Measurement Property Evidence in Adult Populations With Possible Acquired Cognitive Impairments (n=21 studies)

First Author (Year)	Clinical Participant Description (Diagnosis; Mean Age/SD (years); % female)	Measurement Property COSMIN Quality Rating*					
		Internal Consistency	Reliability	Content Validity	Structural Validity	Hypothesis Testing	Responsiveness
Cognitive Competency Test (CCT), n=2							
Rutman[32] (1992)	Not well described, "diverse"; 71/SD not given; 42% Age and sex distribution varied between analyses and not well described	—	—	—	—	P[1,2,4,13]	—
Zur[33] (2013)	Older adult, suspected cognitive impairments; males' age 78.1/13.95 and females' age 81.3/5.62; 60%	†F	—	—	†F	P[14]	—
Everyday Functioning Battery/ Functional Impact Assessment (EFB/FIA), n=3							
Heaton[34] (2004)	Adult, NP impaired and NP normal patients with HIV/AIDS; 39.32/7.52; total sample 16%	P[5]	—	—	—	‡F/P[14]	—
Sadek[35] (2011)	Adult, >6 months poststroke (with or without cognitive impairments); 63.0/10.4; 33%	—	—	—	—	F	—
Harvey[36] (2013) (advanced finances subscale only)	Adult, outpatients with schizophrenia or schizoaffective disorder; ages for 3 recruitment centers: Skyland, 35.78/14.13; Atlanta: 47.3/8.58; and San Diego: 47.25/8.89; the % female for 3 recruitment centers: Skyland, 31%; Atlanta, 25%, and San Diego, 34%	—	F	—	—	—	—
Financial Capacity Instrument (FCI), n=3							
§Marson[37] (2000)	Older adult, community living and with mild or moderate AD; 73.8/7.6; ~62% (discrepancies between numbers and percentages given [p 879])	P[5]	P[2,4,9]	—	—	F	—
‖Marson[38] (2001)	Sample diagnoses and age not described	P[3,5]	P[3,7–9]	—	—	—	—
¶Triebel[39] (2009)	Adult, MCI at baseline (subgroups of people who did and did not convert from MCI to dementia); MCI nonconverters to dementia 68.5/7.5; MCI converters to dementia 74.4/6.0; MCI nonconverters 54.8%; and MCI converters 56%	—	—	—	—	F	F
Financial Competence Assessment Inventory (FCAI), n=2							
Kershaw[40] (2008)	Adult, cognitive impairment (subgroups: ABI, schizophrenia, dementia, and intellectual disability; did not test to ensure all participants had cognitive impairment); 52.68/9.33 (total sample, differences between diagnostic subgroups); 52.8% (total sample, differences between diagnostic subgroups)	P[5]	P[2,4]	—	—	F	—

(*continued on next page*)

Table 2 (*continued*)

First Author (Year)	Clinical Participant Description (Diagnosis; Mean Age/SD (years); % female)	Measurement Property COSMIN Quality Rating*					
		Internal Consistency	Reliability	Content Validity	Structural Validity	Hypothesis Testing	Responsiveness
Pachan[41] (2014)	Adult, community dwelling with dementia or memory impairment; 69.49/9.74; 44%	P[5,6]	—	—	—	F	—
Independent Living Scales (ILS), n=6							
Revheim[42] (2004)	Adult, schizophrenia or schizoaffective disorder; 37.2/8.3; 38%	—	—	—	—	F	—
#Revheim[43] (2006)	Adult, schizophrenia or schizoaffective disorder; 39.2/9.9; 26%	—	—	—	—	P[2,13,14]	—
Baird[44] (2006)	Adult, dementia; 75/6.9; 57.8%	—	—	—	—	F	—
Green[45] (2011)	Adults with schizophrenia; 43.9/10.1; 35.6%	—	**F/ P[2,8]	—	—	P[14]	P[1,4,14]
Velligan[46] (2012)	Bilingual experts from 8 different countries in research field with knowledge of everyday living activities in schizophrenia populations and respective cultures (experts' age and sex distribution not described)	—	—	P[10-12]	—	—	—
‡‡Quickel[47] (2013)	Adult, neurologic referrals for assessment (neurologic inclusion/exclusion not well described, subgroups included people with dementia, schizophrenia, depression, and/or developmental disabilities); 55.6/23.9 (total sample; differences between diagnostic subgroups); 54.9% (total sample)	—	—	—	—	P[3,4]	—
Kohlman Evaluation of Living Skills (KELS), n=3							
Brown[48] (1996)	Adult, consumers of community support program who had persistent mental health issues; 37/SD not given; 45%	n/a	—	—	n/a	P[2,4,14]	—
Pickens[49] (2007)	Adult, identified as having "self-neglect" by caseworkers; 76.3/SD not given; 64%	n/a	—	—	n/a	F	—
Burnett[50] (2009	Adult, identified as having "self-neglect" by caseworkers; 76.5/7.2; 64%	n/a	—	—	n/a	F	—
Measure of Awareness of Financial Skills (MAFS), n=1							
Cramer[51] (2004)	Adult, dementia; 75.6/8.2; 80% For content validity: experts in area of helping people with financial management; expert group age and sex not described)	P[5]	—	P[11]	—	P[2]/F	—

(*continued on next page*)

Table 2 (*continued*)

First Author (Year)	Clinical Participant Description (Diagnosis; Mean Age/SD (years); % female)	Measurement Property COSMIN Quality Rating* Internal Consistency	Reliability	Content Validity	Structural Validity	Hypothesis Testing	Responsiveness
Semi-Structured Clinical Interview for Financial Capacity (SCIFC), n = 1							
Marson[52] (2009)	Adult, amnestic MCI or AD; 68.0/8.3 (MCI group), 72.4/8.4 (mild AD group) and 75.3/8.4 (moderate AD group; differences between diagnostic subgroups); 69% (MCI group), 46% (mild AD group), and 68% (moderate AD group; differences between diagnostic sub-groups)	—	P[1,4,8]	—	—	P[4]	—

Abbreviations: ABI, acquired brain injury; AD, Alzheimer dementia; AIDS, acquired immunodeficiency syndrome; HIV, human immunodeficiency virus; MCI, mild cognitive impairment; NP, neuropsychological; SD, standard deviation; TBI, traumatic brain injury.

* COSMIN quality ratings: P, poor; F, fair; G, good; E, excellent, —, study did not examine this measurement property area (explanation for superscript numbers for poor [P] ratings found in Table 1); n/a, not applicable as instrument is not summed score or is clearly based only on formative measurement model.

† COSMIN rating based on second-order factor analysis and internal consistency (ie, the relation of subscales to total score), not first-order factor analysis and internal consistency (ie, the relation of items to subscales).

‡ Discriminative construct validity (hypothesis testing) evidence quality rated fair/convergent construct validity evidence quality rated poor.

§ FCI-6 (6 subscale structure).

|| FCI-8 (8 subscale structure).

¶ FCI-9 (but only analyzed 8 of the 9 subscales).

ILS-Problem-solving subscale only.

** Test-retest reliability evidence quality rated fair/interrater reliability evidence quality rated poor.

‡‡ ILS money management and home/safety subscales only.

regarding reliability, and 2 had studies regarding responsiveness. However, some MP evidence had ratings of poor methodologic quality and was not considered for adequacy of results and synthesis (table 2). Common reasons for poor ratings of MP evidence were small sample size, dimensionality of the instrument not checked or referenced (for a hypothesized unidimensional scale or subscale), study design issues, and an inadequate description of comparator instrument and its published measurement evidence (see tables 1 and 2).

Regarding diagnostic populations, older adults with mild cognitive impairment, suspected cognitive impairment, or dementia were studied in 9 (of 21) direct MP evidence studies; people living with mental health issues (eg, schizophrenia) were included in 7 studies; people after acquired brain injury (eg, stroke or traumatic brain injury) were included in 2 studies; and mixed participants of the previously mentioned diagnostic groups were included in 2 studies (see table 2).

We outline the MP evidence per instrument in the following text (see table 2 for the methodologic rating of MP evidence per study and appendix 3 for a summary of direct studies included in the final synthesis).

Cognitive Competency Test

We found 2 direct evidence studies. The methodologic quality for construct validity hypothesis testing was poor in both studies. One study included older adults with suspected cognitive impairments but examined only 11 of the 12 CCT subscales ("the Personal Information subscale was not included due to insufficient variability on the item").[33(p.174)] This study found supporting results of fair methodologic quality for second-order structural validity and internal consistency (ie, the relation of the subscales to the total score using factor analysis), indicating that the 11 CCT subscales measure a unidimensional construct.

Everyday Functioning Battery/Functional Impact Assessment

We identified 3 direct evidence studies; one study examined only the advanced finances subscale and not the basic finances subscale.[36] Although all studies identified the same source publication,[34] there were discrepancies in the subscales and scoring structure between studies. For example, Heaton and colleagues[34] and the Functional Impact Assessment (FIA) manual[53] included a vocational assessment subscale, whereas Sadek and colleagues[35] included a communication subscale but no vocational subscale. It remains unclear whether both forms constitute the same instrument; this potential discrepancy presents a challenge when considering total scores aggregated from the subscales. Nonetheless, the 3 studies had fair methodologic quality and the results supported the respective MP evidence areas. The study examining construct convergent validity hypothesis testing reported statistically significant correlations of fair to excellent strength as hypothesized.[35]

Table 3 Synthesis of measurement evidence per instrument in each measurement property area in studies of adult populations with possible acquired cognitive impairments

Instrument	Measurement Property					
	Internal Consistency	Reliability	Content Validity	Structural Validity	Hypothesis Testing	Responsiveness
Cognitive Competency Test (CCT)	*+	—	—	*+	?	—
Everyday Functioning Battery/ Functional Impact Assessment (EFB/FIA)	?	†+ *(test-retest)*	—	—	++	—
Financial Capacity Instrument (FCI)	?	?	—	—	++	CE *(subscale differences)*
Financial Competence Assessment Inventory (FCAI)	?	?	—	—	++	—
Independent Living Scales (ILS)	—	+ *(test-retest)*	?	—	++	?
Kohlman Evaluation of Living Skills (KELS)	n/a	—	—	n/a	++	—
Measure of Awareness of Financial Skills (MAFS)	?	—	?	—	+	—
Semi-Structured Clinical Interview for Financial Capacity (SCIFC)	—	?	—	—	?	—

Abbreviations: +++, strong evidence; ++, moderate evidence; +, limited evidence; CE, conflicting evidence; ?, = unknown evidence due to poor methodologic quality; —, no information available; n/a, not applicable as instrument is not summed score or is clearly based only on formative measurement model.

* Based on second-order factor analysis (ie, relation of subscales to total score), not first-order factor analysis (ie, relation of items to subscales).

† Advanced finance subscale only.

Providing further information for the interpretation of change scores, we used the baseline standard deviation of 3.67 and ICC of 0.74 for the EFB/FIA advanced finance subscale reported by Harvey and colleagues[36] to calculate the SEM_{diff} to be 2.65 ($=3.67 * \sqrt{[2 *(1-.74)]}$) and the SDC_{95} to be 5.19 ($=1.96 * 2.65$).

Financial Capacity Instrument

We found 3 direct evidence studies. The studies of internal consistency and reliability were of poor methodologic quality. Results from 2 studies of construct validity hypothesis testing and 1 study of responsiveness were of fair methodologic quality. Hypothesis testing supported the discriminative construct validity of the FCI. The study of responsiveness had fair methodologic quality but provided mixed support for longitudinal discriminative hypothesis testing.[39]

Financial Competence Assessment Inventory

We identified 2 direct evidence studies. The internal consistency and reliability (ie, interrater, test-retest) results were of poor methodologic quality. Two studies reporting validity were of fair methodologic quality, reporting adequate results to support a priori hypotheses. The study examining convergent validity hypothesis testing reported statistically significant correlations of fair to excellent strength with other instruments related to FM and cognition.[40,41]

Independent Living Scales

We found 6 direct evidence studies. Methodologic quality was fair for 1 study of test-retest reliability evidence and for 2 studies of hypothesis testing construct validity evidence. Test-retest reliability reported adequate ICC results in a population of adults living with schizophrenia.[45] Construct validity evidence from a study of fair methodologic quality reported statistically significant correlations of fair to excellent strength between the ILS and the Dementia Rating Scale.[44]

Using the baseline ILS Total Standard Score SD of 15.1 and ICC of 0.76 reported by Green and colleagues,[45] we calculated the SEM_{diff} to be 10.46 ($=15.1 * \sqrt{[2 * (1-.76)]}$) and the SDC_{95} to be 20.50 ($=1.96 * 10.46$).

Kohlman Evaluation of Living Skills

We identified 3 direct evidence studies. The studies examined construct validity hypothesis testing and only 2 of these studies were of fair methodologic quality. The results from these 2 studies provide evidence for construct validity hypothesis testing.[49,50]

Measure of Awareness of Financial Skills

We found 1 direct evidence study. It examined internal consistency, content validity, and construct validity.[51] The study provides MP evidence of fair methodologic quality for convergent, divergent (discriminant), and discriminative (known groups) construct validity. However, not all hypothesized correlations examining convergent construct validity were supported.

Semi-structured Clinical Interview for Financial Capacity

We found 1 direct evidence study. The study was of poor methodological quality in each MP area studied.

Evidence Synthesis

Amalgamated MP evidence across the instruments is presented in table 3.

Discussion

To our knowledge, this is the first systematic review of MP evidence for FM instruments in which the methodologic quality of

the MP evidence is appraised before appraising the adequacy of results and synthesizing MP evidence. Encouragingly, MP evidence in the area of FM has increased over the past 15 years,[2,13,15] and 5 of the 8 instruments have moderate evidence for construct validity hypothesis testing. However, the overall volume and methodologic quality of the MP evidence for each instrument is low. This may be, in part, because our review examines MP evidence methodologic quality, which is unique to the examination of FM assessment. In other FM-related reviews,[13,15−17] as with many assessment reviews, this crucial step of appraising study quality is lacking. As our results show, MP evidence is often based on studies with a high risk of bias. Instrument recommendations need to be based on MP evidence with a decreased risk of bias.[25]

Considering the dearth of evidence with high methodologic quality and low risk of bias, researchers and clinicians should be cautious when interpreting the results of these 8 instruments for descriptive or evaluative purposes. None of these instruments have comprehensive supporting MP evidence. The FCI, FCAI, and ILS have the largest volume of evidence, although not across all MP areas applicable to how the measures are used (ie, descriptive, evaluative). In congruence with other FM literature, we recommend triangulation of FM assessments and information from multiple sources to obtain more accurate understandings of a person's current FM functions.[2,17,54−56]

Synthesis of MP evidence

Favorably for many instruments, construct validity is supported by the results of direct studies with fair methodologic quality. As shown in appendix 4, 21 studies with indirect MP evidence further support the construct validity for many instruments. Interestingly, the FCI and ILS had the most indirect MP evidence, with 14 and 6 indirect studies, respectively. This may be an indication that they are the most used FM instruments in the published literature; however, use magnitude is not a replacement for MP evidence.

Direct construct validity hypothesis testing evidence was limited to discriminative (known groups) analyses for some instruments, and for many validity studies a priori hypotheses were either unstated or vaguely formulated. When stated, the hypotheses omitted the expected direction and/or magnitudes for correlations or mean differences, which are needed for interpretation of validity evidence.[18,24,25] As such, the results from these studies could only be compared to common correlation strength criteria (ie, weak, moderate, or strong correlation values).

In the MP evidence areas of content validity, reliability, and responsiveness, evidence is unavailable and/or available evidence is of poor methodologic quality. This inadequacy of evidence is worrisome considering that many of these instruments are being used clinically to describe and evaluate groups and individuals.[57] Despite the foundational nature of content validity, we found no study with a methodologic quality rating above poor. Furthermore, these measures are used in populations living with cognitive impairments, but the instruments may not have been designed for individuals with cognitive impairments per se. Hence, the content validity and cognitive accessibility of items may not be optimal; content validity and cognitive accessibility are important to understanding whether poor results are actually a measure of FM limitations versus an issue of the adult being assessed not understanding the items. Without content validity information it is difficult to understand whether the content, layout, or administration procedures are cognitively inaccessible and influencing results.[58]

There are no studies with fair or better methodologic quality examining interrater reliability, an important MP evidence area applicable to research and clinical settings where data from multiple assessors or raters are compared.[18]

Especially concerning is the absence of responsiveness studies, particularly as many of the instruments are reported to have or have been used for an evaluative purpose (eg, FCI, FCAI, ILS). Considering that cognition is a set of dynamic functions that can fluctuate with age, injury or illness, recovery, or in response to medications, the ability to reassess and evaluate the area of FM in populations with cognitive impairments is especially important.[59] Thus, the responsiveness evidence of any instrument measuring FM is essential.[18]

Related to the dearth of responsiveness evidence is the limited acceptable methodologic quality MP evidence for test-retest reliability—namely, the data needed to determine measurement error and interpretation of change scores. Only 2 studies had test-retest reliability data with acceptable quality ratings.[36,45] When the results of these 2 studies were examined, the reliability data met the standards we set for ICC interpretation (ie, $\geq$0.70). However, this standard is for interpreting group data, and it has been suggested that the reliability ICC requirements be more stringent for measurement of individuals (eg, 0.90 or 0.95).[18] This standard was not met by the reported results of these 2 studies.

Of further concern, in individuals who are hypothetically cognitively and/or medically stable, the data indicated relatively large measurement error and SDC_{95} values. A change greater than SDC_{95} is required to be considered to surpass statistical measurement error for reasonable certainty of true change.[18,60] For the ILS full-scale standard score, which has a possible range of 55 to 121,[45,61] we calculated SDC_{95} to be 20.50; for the EFB/FIA advanced finances subscale, which has a possible score range of 0 to 13,[36] we calculated SDC_{95} to be 5.19. Both of these SDC_{95} values are large relative to the size of the respective scales, and large changes are required to be confident that change exceeds measurement error.[18]

Finally, evidence regarding minimal clinically important change is lacking for all instruments. Values of minimal clinically important change provide information about the amount of change needed to be perceived as important or meaningful by the population being assessed, which contextualizes responsiveness evidence and interpretations of SDC for evaluative instruments.[18,62]

Challenges in appraising MP evidence

We encountered 3 main challenges when examining MP evidence. First, some of the instruments have different versions (ie, EFB/FIA, FCI). MP evidence may not be comparable between instrument iterations, especially when a large proportion of items or scales within the instrument are altered.[18] Quality standards require MP evidence to be established for each version. For measures with multiple versions, we indicated in the tables and appendices the associated instrument version specific to certain MP evidence. By appraising specific versions

separately, the volume of MP evidence for some instruments is further reduced.

Second, there were limited FM subscale analyses for instruments in which FM is a subscale (ie, CCT, EFB/FIA, ILS, KELS). Without evidence for the specific subscales related to FM, the reliability and validity of using these subscales alone are uncertain.[18] Only 2 direct evidence studies were of at least fair methodologic quality and specifically examined FM subscales.[36,44]

Third, although the study populations included adults with possible acquired cognitive impairment, the populations across the studies differed by diagnosis and demographic information. Participant demographics used for MP evidence matter— evidence for use in one population may not be generalizable to other populations.[18] In the included studies there is diversity in diagnosis, age, and sex distributions; in 4 studies, participant demographics are not described.[32,38,46,51] We did not find a study of high methodologic quality examining the generalizability of MP evidence across the different cognitive impairment populations. The most represented diagnostic population is one of older adults with mild cognitive impairment, dementia, or suspected cognitive impairment. This is not surprising considering the rising population of older adults, a current cohort tending to have increased assets and elevated rates of fraud and financial abuse perpetrated against them.[8,63] However, other diagnostic populations were included in fewer MP evidence studies (eg, mental health, HIV/AIDS, acquired brain injury), despite the included instruments being used clinically and in research in these populations.

Review limitations

Like all systematic reviews, this review is limited by time and publication biases. It may be that new direct MP evidence literature is unpublished or has been published since our database search. However, it is unlikely that this is the case as our initial search was extensive and our search alerts for related literature have yielded only one additional direct MP evidence. This study of structural validity found that the FCI items fit a 4-factor solution as opposed to the current domain structure of the FCI (in its various iterations).[64] However, this study does not change our conclusion that more high-quality MP evidence is needed for the 8 FM instruments, including the FCI, before clinical or research use. This study actually provides further support of our COSMIN rating of the FCI studies reporting internal consistency (coefficient α) as poor methodologic quality, and the found factor structure raises further questions about the favorable α values reported by these studies of internal consistency (as coefficient α can often be erroneously inflated).[18] A more recent review's search that examined 4 of the measures in this review (ie, FCAI, FCI, MAFS, SCIFC) did not identify new evidence when compared to this review.[15]

Further, although the COSMIN methods have advanced the discussion and consistency of analyzing the methodologic quality of MP evidence, these methods are still evolving.[27,65,66] The appropriateness of the quality rating structure for some MP areas has been questioned, especially when applying the COSMIN tool to observation-based instruments, and some consider the COSMIN standards to be overly complex and stringent.[67,68] However, at the time of this review the COSMIN 4-point rating scale provides the most in-depth and standardized analysis of MP evidence methodologic quality.

Future research

More studies of MP evidence are needed before these instruments can be unequivocally recommended for rehabilitation practice or research, especially in the foundational areas of content validity and reliability.[69,70] Methodologic quality needs to improve in future MP evidence studies and needs to be appraised in assessment instrument reviews, as quality is foundational to the interpretation of MP evidence results.[27,65] Research should also be done on the quality appraisal instruments themselves to assess their validity in assessing flawed or biased MP evidence studies. Work on the COSMIN in underway and appears to be leading to content changes.[64]

Conclusions

We recommend that the 8 instruments be used with caution secondary to the limited methodologic quality of MP evidence and current evidence gaps.[18,60] This is not to say that the instruments are not reliable, valid, or responsive per se; rather, the supporting high-quality MP evidence has yet to be published. All instruments need further empirical support for use as descriptive and/or evaluative instruments.

The assessment and study of FM is developing, and accurate and consistent measurement of FM in adults with cognitive impairment is becoming more critical.[2,7,59] As decisions based on FM instruments can have a significant impact on the well-being and autonomy of adults living with cognitive impairment, FM assessment instruments must have evidence supporting their use and the inferences made based on their results.

Keywords

Adult; Cognition; Financial management; Instruments; Review

Corresponding author

Lisa Engel, PhD, Rehabilitation Sciences Institute, University of Toronto, 500 University Ave, Suite 160, Toronto, ON, M5G 1V7, Canada. *E-mail address:* lisa.engel@mail.utoronto.ca.

Acknowledgment

We thank the librarians at the University of Toronto and Baycrest Health Sciences (Toronto) for assistance in the search protocol development and locating articles.

Appendix 1 Instrument Descriptions and Utility Information

Instrument (Acronym) Country of Development Main References: First Author, (Year)	How to Obtain/ Cost/Time to Administer	Original population developed (Age group; diagnosis)	Score Structure	FM Association	FM structure	Available Versions/Notes Author (year)
Cognitive Competency Test (CCT) Canada Wang[96] (1992)	UO/n.a./∼60 min	Older adults; no specific diagnosis	Combined Average Total Score (ATS), 12 subscale scores	S	U (7 items)	One version in published literature; no longer able to purchase but is still used by some health professions Douglas[57] (2007); Zur[98] (2007)
Everyday Functioning Battery/ Functional Impact Assessment (EFB/ FIA) USA Heaton[34] (2004)	RA/n.a./∼50-60 min for 4 main subscales +∼60 min for optional vocational assessment subscale	Adults; human mmuno-deficiency virus—related neuropsychological deficits	Global Functional Deficit Score (FDS), 7 subscale scores; optional vocational subscale	S	M (3 FM subscales: basic, advanced, all finance	Multiple versions in published literature using different names but all citing Heaton[34] (2004) as source (differences in subscales & scoring); primarily research-based instrument; developed from subscales of available measures and newly developed subscales
Financial Capacity Instrument (FCI) USA Marson[37] (2000)	UO n.a./≥ 60 min Gerstenecker[11] (2016)	Adults; mild cognitive impairment or dementia	1 or 2 summed financial capacity scores (subscale sum scores); 6–9 financial subscale/ "domain" scores (numbers are version dependent)	A	M	Multiple versions: FCI-6 domains (Marson[37], 2000), FCI-7 (Triebel et al.,[92] 2010), FCI-8 (Marson et al[38], 2001; Earnst,[75] 2001), and FCI-9 (Griffith,[77] 2003); currently a research instrument and not marketed to clinical sectors—although in clinical literature has been presented as an instrument option (O'Connor, 2009 has been used in studies of adults with possible acquired cognitive impairments of various diagnoses (eg, dementia, brain injury (Gerstenecker,[76] 2016)
Financial Competence Assessment Inventory (FCAI) Australia Kershaw[40] (2008); Kershaw[80] (2008)	PO/∼$600 CAD/ untimed, ∼30–35 min	Adults; possible cognitive impairments (various diagnoses)	Overall total score, 6 subscales of financial competence, 4 recombined legal competence scores	A	M (total 38 items)	One version in published literature; manual states purpose is both descriptive and evaluative; based on a conceptual model developed in consultation with professionals experienced in helping people with financial issues (Kershaw,[79] 2004)

(*continued on next page*)

Appendix 1 (*continued*)

Instrument (Acronym) Country of Development Main References: First Author, (Year)	How to Obtain/ Cost/Time to Administer	Original population developed (Age group; diagnosis)	Score Structure	FM Association	FM structure	Available Versions/Notes Author (year)
Independent Living Scales (ILS) USA Loeb[61](1996)	PO/ ~$550 CAD/45+ min Bell-McGinty[72] (2002)	Older adults; no specific diagnosis	Full-scale standard score (70 items total); 5 norm-referenced standardized subscale scores of everyday living areas; 2 standardized norm-referenced factor scores (problem solving; performance/ information)	S	U (17 items)	One version in published literature; used by health professions (Douglas,[57] 2007); manual states purpose is both descriptive and evaluative (Loeb,[61] 1996)
Kohlman Evaluation of Living Skills (KELS) USA Kohlman-Thomson[81] (1992)	PO/$140 USD/ ~60–90 min	Adults; mental health issues	17 items in 5 areas; each item ordinal score (independent vs needs assistance)	S	M (6 items)	One version (currently in 4th edition); manual states purpose is descriptive and originally recommended to be part of a comprehensive everyday living skills assessment battery
Measure of Awareness of Financial Skills (MAFS) Canada Cramer[51] (2004) Van Wielingen[93] (2004)	PO/$50 CAD/ 15–20 min for self or proxy reports; ~30–45 min observation-based section	Older adults; dementia	3 analogous parts (self-report, proxy report, observation-based) from which 6 different total scores and comparison scores can be derived. Each part view FM as uni-dimensional	A	U Observation-based scale has 7 items	One version in published literature; developed for use clinically or in research (Van Wielingen,[94] 2004)
Semi-Structured Clinical Interview for Financial Capacity (SCIFC) USA Marson[52] (2009)	UO/n.a./ unknown	Older adults; mild cognitive impairment or dementia	Total score; 9 "domains" scores/ subscales; 3 ordinal rating levels for each domain/ subscale: capable, marginally capable; incapable	A	M	One version in published literature; developed as a clinician-administered semi-structured interview; scoring guidance is provided, but the authors note clinicians "retain autonomy regarding domain and overall capacity judgments" (Marson,[52(p.807)] 2009)

Abbreviations: A, all items (subscales and/or total score) related to financial management or a related construct; FM, financial management; M, multidimensional (multiple FM subscales scores); PO, purchase online; RA, request from author; S, subscale related to financial management or a related construct; U, unidimensional (one FM scale/score); UO, unknown how to obtain.

Appendix 2 Search Terms and Structure Example (Medline/OVID)

Parameters	Search Terms
Population search	1. Cognition Disorders/ 2. Mild Cognitive Impairment/ 3. Brain Injuries/ 4. Brain Diseases/ 5. Schizophrenia/ 6. Nervous System Diseases/ 7. Stroke/ 8. Cerebrovascular Disorders/ 9. Brain Infarction/ 10. Cerebral Infarction/ 11. Cerebral Hemorrhage/ 12. Brain Hemorrhage, Traumatic/ 13. Aged/ 14. Dementia/ 15. Alzheimer Disease/ 16. Lewy Bodies/ 17. Dementia/ 18. Delirium, Dementia, Amnestic, Cognitive Disorders/ 19. AIDS Dementia Complex/ 20. Psychotic Disorders/ 21. Depression/ 22. Depressive Disorder/ 23. Mental Disorders/ 24. Substance-Related Disorders/ 25. Alcoholism/ 26. HIV/ 27. Acquired Immunodeficiency Syndrome/ 28. Brain Damage, Chronic/ 29. Parkinson Disease/ 30. Parkinson Disease/ 31. Multiple Sclerosis/ 32. "acquired brain injur*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 33. "traumatic Brain Injur*".mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 34. ABI.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 35. TBI.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 36. "cerebral vascular accident".mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 37. senile.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 38. senility.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 39. psychosis.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 40. psychiat*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 41. "mental illness".mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 42. "mentally ill".mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 43. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42
Instrument search	44. ("Cognitive Competency Test" or "CCT" or "Everyday Functioning Battery" or "EFB" or "Functional Impact Assessment" or "FIA" or "Financial Capacity Instrument" or "FCI" or "Independent Living Scales" or "ILS" or "Kohlman Evaluation of Living Skills" or "KELS" or "Semi-Structured Clinical Interview for Financial Capacity" or "SCIFC" or "Measure for Awareness of Financial Skills" or "Measure of Awareness of Financial Skills" or "Measure for Assessing Awareness of Financial Skills" or "MAFS" or "Financial Competence Assessment Inventory" or "FCAI").mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
Measurement property search	45. "Reproducibility of Results"/ 46. Observer Variation/ 47. Psychometrics/ 48. Discriminant Analysis/ 49. instrumentation.fs. 50. validation studies.pt. 51. reproducib$.mp. 52. psychometric$.mp. 53. clinimetr$.mp. 54. clinometr$.mp. 55. "observer variation".mp. 56. reliab$.mp. 57. valid$.mp. 58. coefficient.mp. 59. "internal consistency".mp. 60.(cronbach$ and (alpha or alphas)).mp. 61. "item correlation".mp. 62. "item correlations".mp. 63. "item selection".mp. 64. "item selections".mp. 65. "item reduction".mp. 66. "item reductions".mp. 67. agreement.mp. 68. precision.mp. 69. imprecision.mp. 70. "precise values".mp. 71. test-retest.mp. 72. (test and retest).mp. 73. (reliab$ and (test or retest)).mp. 74. stability.mp. 75. interrater.mp. 76. inter-rater.mp. 77. intrarater.mp. 78. intra-rater.mp. 79. intertester.mp. 80. inter-tester.mp. 81. intratester.mp. 82. intra-tester.mp. 83. interobserver.mp. 84. inter-observer.mp. 85. intraobserver.mp. 86. intra-observer.mp. 87. intertechnician.mp. 88. inter-technician.mp. 89. intra-technician.mp. 90. interexaminer.mp. 91. inter-examiner.mp. 92. intraexaminer.mp. 93. intra-examiner.mp. 94. interassay.mp. 95. inter-assay.mp. 96. intraassay.mp. 97. intra-assay.mp. 98. interindividual.mp. 99. inter-individual.mp. 100. intraindividual.mp. 101. intra-individual.mp. 102. interparticipant.mp. 103. inter-participant.mp. 104. intratechnician.mp. 105. intraparticipant.mp. 106. intra-participant.mp. 107. kappa.mp. 108. kappa's.mp. 109. "coefficient of variation".mp. 110. repeatab$.mp. 111. ((replicab$ or repeated) and (measure or measures or findings or result or results or test or tests)).mp. 112. generaliza$.mp. 113. generalisa$.mp. 114. concordance.mp. 115. (intraclass and correlation$).mp. 116. discriminative.mp. 117. "known groups".mp. 118. "factor analysis".mp. 119. "factor analyses".mp. 120. "factor structure".mp. 121. "factor structures".mp. 122. dimensionality.mp. 123.

(*continued on next page*)

Appendix 2 (*continued*)

Parameters	Search Terms
	subscale$.mp. 124. “multitrait scaling analysis”.mp. 125. “multitrait scaling analyses”.mp. 126. “item discriminant”.mp. 127. “interscale correlation”.mp. 128. “interscale correlations”.mp. 129. ((error or errors) and (measure$ or correlat$ or evaluat$ or accuracy or accurate or precision or mean)).mp. 130. “individual variability”.mp. 131. “interval variability”.mp. 132. “rate variability”.mp. 133. “variability analysis”.mp. 134. (uncertainty and (measurement or measuring)).mp. 135. “standard error of measurement”.mp. 136. sensitiv$.mp. 137. responsive$.mp. 138. (limit and detection).mp. 139. “minimal detectable concentration”.mp. 140. interpretab$.mp. 141. (small$ and (real or detectable) and (change or difference)).mp. 142. “meaningful change”.mp. 143. “minimal important change”.mp. 144. “minimal important difference”.mp. 145. “minimally important change”.mp. 146. “minimally important difference”.mp. 147. “minimal detectable change”.mp. 148. “minimal detectable difference”.mp. 149. “minimally detectable change”.mp. 150. “minimally detectable difference”.mp. 151. “minimal real change”.mp. 152. “minimal real difference”.mp. 153. “minimally real change”.mp. 154. “minimally real difference”.mp. 155. “ceiling effect”.mp. 156. “floor effect”.mp. 157. “item response model”.mp. 158. IRT.mp. 159. Rasch.mp. 160. “Differential item functioning”.mp. 161. DIF.mp. 162. “computer adapted testing”.mp. 163. “item bank”.mp. 164. “cross-cultural equivalence”.mp. 165. develop$.mp. 166. utility.mp. 167. or/45-166 168. 43 and 44 and 167 169. addresses.pt. 170. biography.pt. 171. “case reports”.pt. 172. comment.pt. 173. directory.pt. 174. editorial.pt. 175. festschrift.pt. 176. interview.pt. 177. lectures.pt. 178. “legal cases”.pt. 179. legislation.pt. 180. letter.pt. 181. news.pt. 182. “newspaper article”.pt. 183. “patient education handout”.pt. 184. “popular works”.pt. 185. congresses.pt. 186. “consensus development conference”.pt. 187. “consensus development conference, nih”.pt. 188. “practice guideline”.pt 189. or/169-188
Search parameter connections and search filters/limits	190. (“central corneal thickness” or “computed tomography” or “free cortisol index” or “endoscopic forceps biopsy” or “computerized cognitive training”).mp. 191. 168 not (189 or 190) 192. 191 not (animals/ not humans/) 193. limit 192 to “all adult (19 plus years)” 194. limit 193 to english language

Appendix 3 Direct Measurement Property Evidence Results Within Studies of Adult Populations With Possible Acquired Cognitive Impairments

Reference First Author (Year)	COSMIN Rating	MP Evidence	Results
Cognitive Competency Test (CCT)			
Rutman[32] (1992)	Poor	C/K	• Comparison of the CCT-total score and CCT financial management subtest ordinal cutoffs with the MMSE competence found that people deemed competent by a multidisciplinary team tended to score higher on the CCT scores and MMSE. • Some participants who scored "impaired" and "gray area" on CCT (n=4 on full CCT score, n=5 on CCT financial management subscale) were actually found competent by a multidisciplinary team, although on the CCT financial management subscale no one who scored in the impaired range was found competent. No inferential statistics were provided. Participants found incompetent by a multidisciplinary panel tended to have lower CCT scores or incomplete CCT. • Sample size differed between CCT total score and subscale analyses.
Zur[33] (2013) (of 11/12 CCT subscales)	Fair	IC, second order	• Analysis of 11/12 CCT subscales • ++, α 0.826 • n=107
	Fair	S, second order	• Analysis of 11/12 CCT subscales • In factor analysis, one factor had value greater than random generated eigenvalue • n=107
	Poor	C	• Analysis of 11/12 CCT subscales • The CCT score showed significant correlations with the MMSE score, sex, depression, concerns with judgment and insight, and problems found during a kitchen assessment (all $P<.05$).
	Poor	K	• Analysis of 11/12 CCT subscales • MANOVA analysis found significant difference on the CCT between the different levels of the occupational therapy discharge plan variables. • This effect does not show a significant interaction with sex. • The Tukey post hoc analysis CCT score was found to be significantly different (1) between individuals who were discharged to home (with formal support) and individuals who were discharged to long-term care and (2) between individuals who were discharged to a retirement home and individuals who were discharged to long-term care. • Sample size varied between analyses, range n=50–106.
Everyday Functioning Battery/Functional Impact Assessment (EFB/FIA)			
Heaton[34] (2004)	Fair	K	• ++, $P<.0001$(FDS and all subscale scores, neuropsychology deficits HIV group vs neuropsychology healthy HIV group) • n=193–267, differed between subscales.
	Poor	C	• Higher levels of functional impairment on EFB/FIA correlated with a high number of cognitive complaints on the Patient Assessment of Own Functioning Inventory total score; post hoc Pearson r=0.38, $P<.0001$) • The Beck Depression Inventory total score and total EFB-FIA score were not significantly correlated (r=0.10, $P>.05$). • Sample size differed per subscale analysis: cooking n=267; finance n=266, medication management n=193.
Sadek[35] (2011)	Fair	C	• ++, EFB/FIA total score correlated with self ($r=-0.41$, $P<.01$, and Spearman r=−0.51, $P<.001$) and informant (r=−0.48, $P<0.01$, and Spearman $r=-0.54$, $P<.001$) reports using the Pfeffer Functional Activities Questionnaire and the Neuropsychological Assessment Battery total score ($r=0.83$, $P<.001$). • n=45, stroke group only
	Fair	K	• ++, $P<.001$ (total score, healthy controls vs. stroke group) • n=82
Harvey[36] (2013)	Fair	R-TR	• ++, ICC=0.74 ($P=.001$; advanced finance subscale only) • n=195
Financial Capacity Instrument (FCI)			
Marson,[53] (2000)	Fair	K	• ++, on FCI-6, difference in all but one of 6 subscales and all but 5 tasks where people with mild AD performed worse than controls. • Moderate AD group performed worse than mild AD group and healthy controls in all domains and tasks (most $P<.001$, one $P<.05$).

(*continued on next page*)

Appendix 3 (*continued*)

Reference First Author (Year)	COSMIN Rating	MP Evidence	Results
			• ANOVA post hoc analyses found adults with moderate AD performed worse than adults with mild AD and healthy older adult controls on all subscales and tasks; people with mild AD performed statistically significantly worse than healthy older adults controls on all subscales except basic monetary skills (ie, FCI domain 1) and all task areas except naming currency, currency relationships, counting currency, 1-item grocery purchase, understanding checkbook, and detect fraud risk. • n = 73
	Poor	IC	• α range at domain level was 0.85–0.93 and at task level (ie, items within a subdomain) was 0.63–0.89; increased variation at task level. • n = 73
	Poor	R-IR	• Results based on subsample of healthy controls and diagnostic participants rated by 2 independent raters. Analyzed percent agreement for domain and task levels, with more variation at task level. Range of 67% (on task 2a) to 100% (tasks 1a, 1c, 3b, 4b). Except task 2a, all tasks made >80% agreement criterion; domain level ranged 86.4%–100% agreement. • n = 11
	Poor	R-TR	• Results based on subsample of participants with Alzheimer dementia. • Provides Pearson correlations; item-level range 0.78–0.94 and domain-level range 0.85–0.98. • All test-retest coefficients were significant ($P<.001$) except task 6a ($P<.05$, $r=0.50$) • n = 17
Marson[38] (2001)	Poor	IC	• Derivation of results were not well described. • α varied between domain and task levels. Domains had higher α's overall (domains 1–7 α range 0.81–0.93; domain 8 α 0.41). Task-level α range 0.10–0.95.
	Poor	R-IR	• Derivation of results were not well described. • Provides percent agreement only; reported range 90%-100% • Sample size for interrater reliability: domains 1–7, n = 40; domain 8, n = 35
	Poor	R-TR	• Derivation of results were not well described • Provides Pearson correlations; variance. • In correlations between different task levels (range 0.46–0.95); less variance at domain level (range 0.78-0.92). Sample and reliability methods not explained in the article. • Sample size for test-retest varies: domains 1–7, n = 42; domain 8, n = 30.
Triebel[39] (2009)	Fair	K	• ++, on FCI-9 (although no analysis of subscale 8) adults with MCI vs healthy controls at baseline in 3 FCI subscales and 2 FCI summed scores (all $P<.002$) • n = 163
	Fair	Re	• Mixed results. Discriminative change between people with MCI at baseline who converted to dementia after 1 year vs MCI at baseline who did not convert to dementia; only 3 FCI scores statistically significant ($P<.002$) • n = 163
Financial Competence Assessment Inventory (FCAI)			
Kershaw[40] 2008	Fair	C	• ++, FCAI statistically significantly correlated with tests related to financial management and cognition: the Hopemont Capacity Assessment Interview financial decision-making scale (total sample analysis, total FCAI score ($r=0.85$, $P<.01$); subscales r = 0.40–0.82 (range across 6 subscales, all $P<.01$); the Independent Living Scales money management subscale (total sample analysis, total FACI score $r=0.89$, $P<.01$); subscales $r=0.45$–0.87 (range across 6 subscales, all $P<.01$); and the MMSE (total sample analysis, total FCAI score r = 0.85, $P<0.01$; subscales $r=0.47$–0.88 (range across 6 subscales, all $P<.01$) • n = 178 (adults with probable cognitive impairment of various diagnoses)
	Fair	K	• ++, FCAI total score and all 6 subscales were able to statistically significantly discriminate between people with and without a legally appointed financial administrator/guardian; adults without an administrator performed better on the FCAI (total FCAI score comparison, Wilks λ: F(7167) = 22.15, $P<.01$, partial q = 0.486). • Subscale analysis independent *t* tests with Bonferroni-type adjustment (all $P<.01$) • n = 178

(*continued on next page*)

Appendix 3 (*continued*)

Reference First Author (Year)	COSMIN Rating	MP Evidence	Results
	Poor	IC	• α for 38-item full-scale total α=0.96. Five scales had good internal consistency (everyday financial abilities α =0.89; financial judgment α =0.86; estate management α =0.84; cognitive functions regarding financial tasks,α=0.91); debt management α =0.84). • One scale had low internal consistency (support resources α =0.54). • n=178 in a mixed sample
	Poor	R-IR	• Percent agreement range 83%–98% on all 38 items • Cohen κ for all-item agreement 0.86 • n=10
	Poor	R-TR	• Pearson correlations between 2 measurement points (range 5–9 weeks apart; n=20 healthy adult comparison groups): total scale (*r*=0.93), debt management (*r*=0.57), cognitive functions related to financial tasks (*r*=0.69), everyday financial abilities (*r*=0.97), financial judgment (*r*= 0.95), estate management (*r*=0.93), and support resources(*r*=0.98) • n=20 (subsample of control/healthy group only)
Pachana[41] (2014)	Fair	K	• ++, older adults with cognitive impairments/dementia performed worse on the FCAI total score (without the debt management subscale) than healthy older adult controls (Kruskal-Wallis χ^2=28.9, *df*=1, *P*=.0001) • Logistic regression model for group membership (eg, healthy controls vs dementia) prediction by FCAI subscale analysis found only 2 subscales statistically significant: scores on the Cognitive Functioning subscale (odds ratios=0.42; 95% CI 0.21–0.82) and Support Resources subscale were independently associated with group membership, with (odds ratio=0.73, 95% CI 0.57–0.93). • n=97
	Poor	IC	• FCAI full-scale, including the debt management subscale (FCAI-a): α=0.87; FCAI scale without the debt management subscale (FCAI-b): α=0.91. • Sample size for IC analysis was not well described; assumed clinical and controls combined (n=101).
Independent Living Scales (ILS)			
Revheim[42] (2004)	Fair	K	• ++, ILS problem-solving scores differed among people needing maximum daily supervision, moderate supervision, and minimum supervision (ANOVA *P*<.01, n=162; Bonferroni post hoc analysis differences between all groups, *P*<.002) • n=162
Revheim[43] (2006)	Poor	C	• Problem-solving subscale (ILS-PB) only; no a priori hypotheses stated • Clinical symptoms significantly correlated with the ILS-PB include the Brief Psychiatric Rating Scale withdrawal-retardation factor (*P*=.002); Scale for the Assessment of Negative Symptoms (SANS) total score (*P*<.001); SANS Affective Flattening (*P*=0.001); and Alogia global items (*P*=.005, Bonferroni correction). • Multiple neurocognitive measures significantly correlated with the ILS-PB (Bonferroni corrections; all *P*<.01). Neurocognitive measures were not significantly correlated with ILS-PB including Grooved Pegboard, Brief Visual Memory Test, Categorical Verbal Fluency, Wechsler Adult Intelligence Scale–III, Perceptual Organization Index, and Wisconsin Card Sorting Test (Bonferroni correction, all *P*>.01) • Sample sizes differed for some analyses; range n=30–38.
	Poor	K	• Problem-solving subscale (ILS-PB) only; no a priori hypotheses stated • ILS-PB was significantly different between inpatient and outpatient participants with schizophrenia (ANOVA; *F*=8.7, *P*<.006). • Sample sizes differed for some analyses; range n=30–38.
Baird[44] (2006)	Fair	C	• ++, ILS Full-Scale Standard Score correlated with Dementia Rating Scale score (*r*=0.80, *P*<.05) and Geriatric Depression Scale (*r*=−0.32, *P*<.05). ILS Money Management subscale correlated with DRS subscales (*r*=0.28–0.63, *P*<.05). • n=83
	Fair	K	• ++, ILS Full-Scale Standard Score group means differ between older adults with normal vs impaired Dementia Rating Scale scores (*P*<.05, n=83)

(*continued on next page*)

Appendix 3 (*continued*)

Reference First Author (Year)	COSMIN Rating	MP Evidence	Results
Green[45] (2011)	Fair	R-TR	• ++, Total ILS ICC=0.76, ILS performance factor score ICC=0.70, ILS problem-solving factor score ICC=0.68 ($P<.001$; test-retest; n=144)
	Poor	R-IR	• ICC=0.73 • Videotapes of assessment were evaluated by 3 raters. • n=8
	Poor	C	• MATRICS Consensus Cognitive Battery as measure of cognitive performance was significantly correlated with ILS full-scale score (r=0.51), performance subscale (r=0.53), and problem-solving subscale (r=0.39). • Quality of Life Scale was significantly correlated with ILS full-scale score (r=0.30), performance subscale (r=0.24), and problem-solving subscale (r=0.27). • n=16
	Poor	Re	• Baseline to 4-week follow-up repeated-measure significant result but with small effect size 0.15 (ILS full-scale), 0.10 (ILS performance subscale), and 0.07 (ILS problem solving subscale) • An intervention was not given between repeated measure; therefore interpretation may be an influence of practice effects. • n=144
Velligan[46] (2012)	Poor	CV	• The ILS was rated as being less adaptable to rural residents in China, India, and Mexico. • The ILS was rated as more sensitive to urban/rural differences in India and Mexico than in the United States. • The ILS Money Subscale failed to reach the content validity criterion score in every country, and the Home subscale failed to meet the criterion in India. • n=55–56; expert raters with discrepancy in sample sizes reported in text and table
Quickel[47](2013)	Poor	K	• Money management (ILS-mm) and health and safety (ILS-hs) subscales only • Diagnostic groups differed significantly on the ILS-mm scale ($F(3, 24)=3.00$, $P=0.05$, partial $\eta^2=0.27$) • Tukey post hoc test found people with neurologic diagnoses performed better on the ILS-mm ($P=.06$) and the ILS-HS ($P=.10$) compared with individuals with intellectual disabilities or developmental delays. • People who were found competent by the courts had statistically significantly higher scores than people found incompetent. However, court determination of competence was partly based on the reported ILS measure reported and determination could have been influenced by ILS results. • n=47 for ILS-mm and n=48 for ILS-hs
Kohlman Evaluation of Living Skills (KELS)			
Brown[48] (1996)	Poor	K	• Hypothesized that persons with severe mental illness would perform better on the KELS (interview and simulated task instrument) than in the natural environment because of the additional complexities presented by performance in context; the hypothesis was partially supported. • There were subtask differences; specifically for money subtask there was significant correlation ($\chi^2=5.96$, $P<.05$). • Descriptive analysis: 17/20 (85%) independent in KELS and naturalistic test; 2/20 independent KELS, but assist in naturalistic; 1/20 assist in both • n=20
Pickens[49](2007)	Fair	K	• ++, Self-neglect group had significantly increased total fails on KELS (fail was a >5/16 on KELS) versus healthy matched controls ($\chi^2=5.0$, $P=.025$) • n=92
Burnett[50] (2009)	Fair	C	• ++, The KELS score was correlated with measures of physical, cognitive, and affective function in a group of adults referred to protective services for self-neglect (6 different measures, r=+ to −0.32–0.77, all $P<.013$) • The KELS score was most strongly correlated with the Executive Interview (r=0.77, $P<.001$). KELS was not significantly correlated in this group with the 8-foot walk test and the Geriatric Depression Scale • n=92

(*continued on next page*)

Appendix 3 (*continued*)

Reference First Author (Year)	COSMIN Rating	MP Evidence	Results
	Fair	K	• ++, Adults with self-neglect had lower KELS scores than healthy matched community-living control group (*P*<.001). • n=200
Measure of Awareness of Financial Skills (MAFS)			
Cramer[51] (2004)	Fair	C	• ++, Statistically significant correlations between the MAFS and the Modified MMSE cognitive status measure (*r*=−0.41, *P*<.004, n=35) • No statistically significant correlations between MAFS and physician ratings of FM awareness (*r*= −0.051, *P*=.444, n=10) and Mastery Score (*r*=0.008, *P*=0.486) • n=35
	Fair	D	• ++, No statistically significant correlations between the MAFS and NEO-N Personality Inventory negative emotions measure (*r*=0.31, *P*=.073, *n*=35); and Marlowe-Crowne Social Disability scale (*r*=0.045, *P*=.799) • n=35
	Fair	K	• ++, Older adults with dementia scored statistically significantly worse than healthy older adult controls on MAFS self-report score (*P*=.011), MAFS informant report score (*P*<.001), Observation-based score (*P*<.001) and MAFS unawareness score (*P*<.001) • n=35
	Poor	IC	• Self-report α=0.92; proxy/informant report α-0.97. Performance scale was not analyzed. • n=35
	Poor	CV	• All experts agreed that 18/19 (95%) of objectives were represented by the items, although there was disagreement as to which objectives were linked to which items. • All experts agreed that there was no item about the objective “who provides them with help” • 14/34 items in the questionnaire had 100% agreement (6/6 experts) as to which objective they referred to; 19/34 items had 5/6 experts agree on item to objective match in questionnaire; 5/6 experts agreed item to objective match on 5/6 of the performance item • n=6 experts
	Poor	C	• Completed using the MAFS “unawareness score” (informant questionnaire total score minus self-report questionnaire total score). Performance scale of MAFS not part of convergent or divergent analysis. • MAFS and physician ratings of financial management awareness showed no correlation (*r*=−0.051 *P*=.444); MAFS and Modified MMSE were correlated (*r*=−0.42, *P*=.004); MAFS and Mastery Score was not significantly correlated (*r*=0.008, *P*=.486).
	Poor	D	• MAFS and Five-Factor Inventory Neuroticism Scale to test for negative emotions not correlated as hypothesized (*r*=0.307, *P*=.073 (although noted trend to correlation) • MAFS and Marlowe-Crowne Social Desirability scale not-correlated as hypothesized (*r*=0.045, *P*=.799) • Sample size differed hypothesis testing analyses; range n=10−35
	Poor	K	• Significant differences between groups on t test between well-senior group and dementia group on participant total questionnaire (*P*=.011, informant total questionnaire (*P*<.001), performance total (*P*<.001), and unawareness score (*P*<.001). • Sample size differed in hypothesis testing analyses; range n=10−35
Semi-Structured Clinical Interview for Financial Capacity (SCIFC)			
Marson[52] (2009)	Poor	R-IR	• Percent of participants that 5/5 raters agreed and 4/5 agreed listed in table for 8 domains and domains 1−7 total given for total group analysis and 4 diagnostic subgroups • Large range of participant percentages to meet 80% agreement (4/5 physician agree) criterion (eg, 39% of participants had 4/5 physicians agree on ratings for assets and estate management in 1 diagnostic subgroup vs some ratings in some diagnostic groups had 4/5 physicians agree for 100% of the participants)

(*continued on next page*)

Appendix 3 (*continued*)

Reference First Author (Year)	COSMIN Rating	MP Evidence	Results
	Poor	K	• Differences among healthy controls, participants with MCI, mild AD, and moderate AD. • Between-group differences ($P<.01$) were found for all 8 domains of SCIFC and overall total capacity SCIFC ordinal rating. Post hoc analysis of impairment level: controls<MCI in 7/9 ratings (not bank statement and overall rating); controls and MCI<mild and moderate AD on 7/9 ratings (not basic money management or overall rating); controls<moderate AD on all 9 ratings; mild AD<moderate AD on 7/9 ratings (not checkbook management or bill payment). • n=261 (mixed controls and diagnostic groups)

NOTE: Participant demographics are described in Appendix 4.

++, Results deemed adequate per standards outlined in methods (for fair- or better-quality evidence only); α, Cronbach's α; AD, Alzheimer dementia; ANOVA, analysis of variance; C, convergent construct validity (hypothesis testing); CI, confidence interval; CV, content validity; D, divergent/discriminant construct validity (hypothesis testing); DRS, Dementia Rating Scale; FDS, Functional Deficit Score; FM, financial management; HIV, human immunodeficiency virus; IC, internal consistency; ICC, intraclass correlation coefficient; K, discriminative/known-groups construct validity (hypothesis testing); MANOVA, multivariate analysis of variance; MCI, mild cognitive impairments; MMSE, Mini-Mental State Examination; MP, measurement property; *P*, statistical significance level; *r*, Pearson (parametric) or Spearman (nonparametric) correlations; Re, responsiveness/sensitivity to change hypothesis testing; R-IR, interrater reliability; R-TR, test-retest reliability; S, structural construct validity.

Appendix 4 Indirect Measurement Property Evidence Results in Studies of Adult Populations With Possible Acquired Cognitive Impairments (n=21 Studies)

First author (Year)	Clinical Participant Description (Diagnosis; Mean Age in Years/SD; % female)	MP Evidence	Results
Cognitive Competency Test (CCT), n=1			
Christensen[73] (2005)	Adults with schizophrenia 30.8/SD not given; sex distribution not described	**C**	• Measure of insight/awareness of current and past mental illness had positive, moderate/good strength correlations with CCT (*r*=0.50, *P*< .001; ie, higher insight had higher cognitive competency) • n=40
		K	• Participants who measured as having more disorganized thoughts obtained statistically significant lower CCT scores. • n=40
Everyday Functioning Battery/Functional Impact Assessment (EFB/FIA), n=0			
No indirect psychometric evidence studies found in search			
Financial Capacity Instrument (FCI), n=14			
Earnst[75] (2001)*(c)	Older adults with probable mild or moderate AD; 71.9/7.2; sex distribution not described	C	• Six FCI subscales and FCI Total Score (summed subscales 1–7) had statistically significant moderate to strong correlations with measures of executive function (Wechsler Adult Intelligence Scale-III Digits Backward, Arithmetic, and Letter-Number Sequencing test; all *r*>0.50, *P* <0.05) • n=20 with AD
		D	• Mixed results but mostly supported divergent hypothesis • Two subscale task areas (tipping, understanding a bank statement) showed statistically significant moderate correlations with measure of phonological loop (ie, Wechsler Adult Intelligence Scale-III Digits Forward test; *r*>0.40, *P*<.05) • n=20 with AD
		K	• All 8 domains and all tasks had statistically significant differences between controls and older adults with AD, with most *P*=0.001 and all *P*<.021; summed score of domains 1–7 statistically significant difference *P*=.001 • n=23 healthy controls and n=20 older adults with AD
Griffith[77] (2003)*(d)	Adults with amnestic MCI and mild AD; 68.1/8.8 (MCI group) and 71.5/9.2 (mild AD group); 52.4% (MCI group) and 59.1% (mild AD group)	K	• Healthy controls and people with MCI performed statistically significantly better than subjects with mild AD on the total summed score (subscales 1–8) and all domains with the exception of Financial Judgment (subscale 6) and Knowledge of Personal Assets/Estate (subscale 8) • Healthy controls performed significantly better than the MCI group on Financial Concepts (domain 2), Bank Statement Management (domain 5), and Bill Payment (domain 7) • All *P*<.05 with many *P*=.001 • n=15–21 healthy controls, n=13–21 people with MCI, and n=18–22 people with mild AD (sample size varied between analyses)
Wadley[95] (2003)*(c)	Older adults with AD; 71.9/7.2 (AD group); 50% (AD group)	C	• Convergent validity of FCI with questionnaire-based reports of current FM had mixed results • Discrepancies between performance scores of FCI-8 subscales/summed score of FCI and self and family ratings using the current Financial Capacity Form (n=20 older adults with AD and a family caregiver dyad) • More or less discrepancies dependent on FCI subscale and type of questionnaire rating (self vs informant)
Okonkwo[88] (2006)*(e)	MCI; 69.54/8.22; 44%	C	• Attention measures (Dementia Rating Scale-Attention; Wechsler Memory Scale-spatial span forward) correlated with 2 FCI subscales: Financial Conceptual Knowledge (coefficient of determination=0.26, *P*=.001) and Bank Statement Management (coefficient of determination=0.22, *P*=.001) • Executive function measures (Dementia rating scale-initiation/perseveration; Trail Making Test-part A, Executive Clock Drawing Task; Wechsler Adult Intelligence Scale-digit symbol) correlated with FCI subscale Bill Payment (coefficient of determination=0.34, *P*=.001) • n=43
		K	• Those with MCI performed statistically significantly worse than healthy controls on FCI domains of financial conceptual knowledge, bank statement

(*continued on next page*)

Appendix 4 (*continued*)

First author (Year)	Clinical Participant Description (Diagnosis; Mean Age in Years/SD; % female)	MP Evidence	Results
			management, and bill payment ($P<.008$), but not on cash transactions ($P=.101$). • MCI group needed more time to complete multistep FM task. • n=43 MCI and n=43 healthy controls
Martin[82] (2008)*(d)	Mild AD, community living; 70.6/8.4; 44	K	• Statistically significant differences between those with AD and healthy control groups in age and sex • Mild AD group performed worse than healthy older adult controls on 2 total summed scores ($P=.001$), 9 subscale scores ($P<.009$ with most $P=.001$), and 16 of the 18 financial capacity tasks ($P<.009$ with most $P=.001$) • Mild AD group performed equivalently with the control group on 2 tasks: naming coins/currency ($P=.16$) and a 1-item grocery store transaction ($P=.03$) • n=55 mild AD and n=63 healthy older adult controls
		Re	• Statistically significant differences between AD and healthy control groups in age and sex • Using paired-samples *t* test, the mild AD group had within-group 1-year declines on 2 summed total scores ($P=.001$), 7 of 8 subscales ($P<.007$), except Knowledge of Assets subscale, ($P=.72$), and 12 of 18 tasks ($P<.01$). • Healthy older adult controls had no statistically significant change across any of the 2 total summed scores, 9 subscales, or 18 tasks relative to baseline ($P>.03$). • ANCOVA analysis found group-by-time interaction was statistically significant for 2 total summed scores ($P=.001$), 4 subscale scores ($P<.005$; Basic Monetary Skills, Cash Transactions, Bank Statement, and Investment Decision Making), and 5 of 18 tasks ($P<.01$). • n=55 mild AD and n=63 healthy older adult controls.
Okonkwo[87] (2009)*(e)	MCI (70.1/8.1; 47%)	K	• MCI group had significantly more difficulties than healthy control group on summed score of 4 FCI domains (financial conceptual knowledge, cash transactions, bank statement management, bill payment; $P=.001$) • n=57 MCI and n=68 healthy controls
Sherod[90] (2009)*(d)	MCI and mild AD; 70.3/7.4 (MCI group) and 73.8/8.5 (mild AD group); 57% (MCI group) and 44% (mild AD group)	K	• Statistically significant differences between diagnostic groups and healthy controls in age • The FCI summed score (subscales 1–7) differed across the 3 study groups ($P<.001$). • Post hoc analyses found that healthy control group performed statistically significantly better than MCI group and mild AD group; MCI patients performed significantly better than the patients with mild AD. • n=113 MCI, n=48 mild AD, and n=85 healthy controls
Griffith[78] (2010)*(e)	Amnestic MCI; 70.8/6.4; 66%)	C	• In amnestic MCI group, FCI had statistically significant, moderate strength correlations with Wechsler Memory Scale-spatial span subtests for attention ($r=0.58$, $d=1.42$, $P=.01$), and the Trail Making Test—part A for visuomotor tracking ($r=-0.56$, $d=1.35$, $P=0.01$); excellent strength correlations with the Wide Range Achievement Test-Arithmetic subtest for arithmetic ability ($r=0.74$, $d=2.20$, $P=.01$). • n=38 amnestic MCI group
		K	• The healthy control group performed significantly better than people with amnestic MCI ($P=.001$, $d=1.38$) on summed FCI score. • n=38 MCI and n=28 healthy controls
Triebel[92] (2010)*(b)	Amnestic MCI; 71.1/7.3 (white participants) and 68.4/8.2 (African American participants); 48.3% (white participants) and 78.9%	K	• Statistically significant differences between cultural/racial subgroups in age and gender • People with amnestic MCI and white cultural/racial heritage had statistically significantly higher scores on FCI summed score and 6 FCI subscale scores ($P<.005$) than people with amnestic MCI and African American cultural/racial heritage.

(*continued on next page*)

Appendix 4 (*continued*)

First author (Year)	Clinical Participant Description (Diagnosis; Mean Age in Years/SD; % female)	MP Evidence	Results
	(African American participants)		• No statically significant difference between groups on Financial Judgment (subscale 6; $P=.159$). • n=112–116 amnestic MCI and white cultural/racial heritage, and n=36–38 amnestic MCI and African American cultural/racial heritage (sample size varied between analyses)
Dreer[74] (2012)*(d)	TBI, moderate to severe, inpatients (exact time from injury not given, ~30 d after end of injury-related confusion/amnesia cleared); 30.0/11.7; 33% (same sample as Martin et al[20])	K	• People with TBI at baseline had statistically significant worse FCI scores compared with healthy adults controls on 2 total summed scores ($P=.001$) and all 9 FCI subscales ($P=.001$) except Basic Monetary Skills ($P=.12$). • At 6-month follow-up the TBI group still had worse FCI scores compared with the healthy adult control group on 2 summed total scores ($P=.001$) and 6 of 9 subscale scores ($P<.29$). • n=24 TBI and n=20 healthy adult controls
		Re	• Group (TBI vs healthy controls) by time (baseline vs 6-month follow-up) MANOVA results demonstrate statistically significant interactions for 2 summed total FCI scores ($P=.001$) and 4 of 9 FCI subscales: Basic Monetary Skills ($P=.26$), Cash Transactions ($P=.001$), Bank Statement Management ($P=.001$), and Bill Payment ($P=.001$). • n=24 TBI and n=20 healthy adult controls
Martin[83] (2012)*(d)	TBI, moderate to severe, inpatients (exact time from injury not given, ~30 d after end of injury-related confusion/amnesia cleared); 30.0/11.7; 33% (same sample as Dreer et al[74])	C	• In the TBI group, at baseline the FCI had statistically significant correlations of good to excellent strength with Wechsler Adult Intelligence Scale-III-Arithmetic subscale ($r=0.79$; $P=0.001$), Wechsler Memory Scale-logical memory subscale ($r=0.59$, $P=.01$), and the Tokens Test ($r=0.61$, $P=.002$), all $P<.05$). • At 6-month follow-up, multiple cognitive measures, particularly executive functions measures, were strongly correlated with FCI total scores: Tokens Test ($r=0.74$, $P=.001$), Wechsler Adult Intelligence Scale-III-Arithmetic subscale ($r=0.70$, $P=.001$), Trail Making Test-part B ($r=-0.66$, 0.001). • n=24
		K	• People with TBI had statistically significant worse FCI total summed scores (subscales 1–7) compared with healthy adults controls at baseline ($P=.001$) and at 6-month follow-up ($P=.001$). • n=24 TBI and n=20 healthy adult controls
		Re	• Group (TBI vs healthy controls) by time (6-month follow-up) MANOVA results demonstrate statistically significant interactions for summed total FCI scores ($P=0.001$). • n=24 TBI and n=20 healthy adult controls
Martin[84] (2013)*(d)	PD with MCI or dementia; 66.9/9.3 (PD MCI group; 71.0/6.1 (PD dementia group); female 50% (PD MCI group); 23.5% (PD dementia group)	K	• PD-related dementia group performed significantly worse than healthy control and PD-MCI groups on 2 FCI summed scores and all FCI domains ($P<.001$) except Financial Judgment (subscale 6) and Knowledge of Personal Assets/Estate Arrangements (subscale 8). • The PD-MCI group performed statistically significantly worse than healthy control group on 2 FCI summed scores and on 3 subscales (Basic Monetary Skills, Financial Concepts, and Investment Decision Making; $P<.001$). • n=17 PD dementia, n=18 PD-MCI, and n=20 healthy controls
Stoeckel[91] (2013)*(d)	Mild AD, community living; 76.9/8.2; 35.2%	C	• Adults with mild AD, FCI Total score was correlated with the Dementia Rating Scale-2 Total score, and the Dementia Rating Scale subscales of Attention, Construction, and Conceptualization (all $P<.005$). • n=16
		K	• On the abbreviated FCI Total score, the mild AD group performed statistically significantly worse than healthy control group ($P=.001$). • n=16 mild AD and n=16 healthy controls
Financial Competence Assessment Inventory (FCAI), n=0			
No indirect psychometric evidence studies found in search			

(*continued on next page*)

Appendix 4 (*continued*)

First author (Year)	Clinical Participant Description (Diagnosis; Mean Age in Years/SD; % female)	MP Evidence	Results
Independent Living Scales (ILS), n=6			
Baird[71] (2001)	Older adults referred for neuropsychologic assessment; 74/SD not given; 56.5%	C	• ILS Full Scale score, 2 factor scores, and 4 of 5 subscale scores were statically significantly correlated ($P<.05$) with 10 different neuropsychological measures (e.g., Dementia Rating Scale, Trail Making Test-parts A and B, Wide Range Achievement Test). • Statically significant correlations ranged from weak (eg, $r=0.20$ between Tokens Test and ILS Full Scale Score) to moderate/strong (eg, $r=0.75$ between Performance factor score and the Dementia Rating Scale). The ILS Social Adjustment subscale was only statistically significantly correlated with the Geriatric Depression Scale ($r=-0.53$, $P<.05$). • n=69 older adults referred for clinical neuropsychological assessment
Bell[72] (2002)	Older adults referred for assessment regarding possible dementia diagnosis; 74.5/6.4; 62.0%	C	• ILS Full Scale Standard Score had statistically significant correlations of fair to excellent strength with tests of executive function (all $P<.001$): Trail Making Test-part B ($r=-0.71$), Wisconsin Card Sorting Test ($r=-0.53$), Dementia Rating Scale Initiation/Perseveration Index ($r=0.44$), Controlled Oral Word Association Test ($r=0.53$), and Manual Postures Test ($r=-0.49$) • Subscales had mixed correlation results with previous listed tests of executive functions; not all had statistically significant correlations, strength of statistically significant correlations varied. • The ILS Money Management Scale had statistically significant correlations of fair to good/moderate strength with 4 tests of executive function (all $P<.001$): Trail Making Test-part B ($r=-0.65$), Wisconsin Card Sorting Test ($r=-0.51$), Controlled Oral Word Association Test ($r=0.49$), and Manual Postures Test ($r=-0.45$) • The Dementia Rating Scale Initiation/Perseveration Index was not statically significant correlated with ILS Money Management Scale, • n=50
Revheim[89] (2004[†])	Schizophrenia or schizoaffective disorder, inpatients and outpatients; 37.2/8.3; 37.7%	C	• ILS problem-solving factor score had statistically significant correlations of fair to moderate/good strength with tests of verbal memory and problem solving (all $P<.01$): Wechsler Adult Intelligence Scale-comprehension ($r=0.63$), Wechsler Memory Scale-logical memory/immediate recall ($r=0.65$), California Verbal Learning Test-list recall ($r=0.62$), California Verbal Learning Test-conceptual clustering ($r=0.40$), and California Verbal Learning Test-consistency ($r=0.39$). • The California Verbal Learning Test-serial order subscale was not statistically significantly correlated. • n=162 (n=87 inpatients and n=75 outpatients)
		K	• Statistically significant differences found for ILS problem-solving factor score ($P<.01$). • The scores of inpatients were significantly lower than the scores of outpatients. • n=87 inpatients and n=75 outpatients
Weiner[97] (2006)	Dementia with mild to moderate cognitive impairment; 71.2/8.2; 26%	C	• In this study the ILS was used as the criterion measure; therefore indirect evidence. • Total ILS score and 4 ILS subscale scores (ie, memory/orientation, money management, home and transportation, and health and safety) were statistically significantly correlated with the Test of Everyday Functional Abilities total scores (all $P<.001$; correlations ranged from moderate ($r=0.62$) to strong ($r=0.89$). • The Test of Everyday Functional Abilities memory subscale did not statistically significantly correlate with Total ILS score or any of the 4 ILS subscale scores analyzed • n=27

(*continued on next page*)

Appendix 4 (*continued*)

First author (Year)	Clinical Participant Description (Diagnosis; Mean Age in Years/SD; % female)	MP Evidence	Results
Medalia[85] (2008[†])	Schizophrenia with cognitive impairment; 38.9/11.4; 26.8%	C	• Correlations and agreement analysis (ie, κ) between the ILS problem-solving factor score and Measure of Insight into Cognition self-report were not statistically significant, indicating poor agreement between subjective and objective measures of cognition related to everyday living. • n=71
Mills[86] (2014)	Older adults referred for outpatient capacity evaluation; 76/10.9; 57.1%	C	• In this study the ILS was used as the criterion measure; therefore indirect evidence. • ILS Full Scale Standard Score had statistically significant correlation of fair strength with the Making and Executing Decisions for Safe and Independent Living test ($r=0.44$, $P<.01$). • n=49
Kohlman Evaluation of Living Skills (KELS), n=0			
No indirect psychometric evidence studies found in search			
Measure of Awareness of Financial Skills (MAFS), n=1			
Van Wielingen[93] (2004)	Adults with varying subtypes of dementia; 76.9/6.6; 57.1%	C	• Weak negative correlation was found between the level of financial skills awareness on the MAFS and dementia severity on the Modified MMSE ($r=-0.284$, $P<.05$). • Moderate positive correlation was found between the MAFS and the Frontal Behavioural Inventory (measure of executive functioning; $r=0.414$, $P<.05$, 1-tailed). • n= 42
Semi-Structured Clinical Interview for Financial Capacity (SCIFC), n=0			
No indirect psychometric evidence studies found in search			

Abbreviations: AD, Alzheimer's dementia; AIDS, acquired Immunodeficiency syndrome; ANCOVA, analysis of covariance; C, Convergent construct validity (hypothesis testing); d, effect size; D, divergent/discriminant construct validity (hypothesis testing); FM, financial management; HIV, human immunodeficiency virus; K, discriminative/known-groups construct validity (hypothesis testing); MCI, mild cognitive impairment; MP, measurement property; PD, Parkinson disease; Re, responsiveness /sensitivity to change hypothesis testing; SD, standard deviation; TBI, traumatic brain injury.

* (a) FCI-6; (b) FCI-7; (c) FCI-8; (d) FCI-9; (e) 4 selected subscales(subscales differ between studies).

† ILS-Problem-solving subscale only.

References

1. American Bar Association Commission on Law and Aging/American Psychological Association Assessment of Capacity in Older Adults Project Working Group. Assessment of older adults with diminished capacity: a handbook for psychologists. Chicago: American Bar Association; 2008. Available at: http://www.bit.ly/2ttW1mn. Accessed May 5, 2015.
2. National Academies of Sciences, Engineering, and Medicine. Informing Social Security's process for financial capability determination. Washington, DC: National Academies Press; 2016.
3. Corrigan JD, Whiteneck G, Mellick D. Perceived needs following traumatic brain injury. J Head Trauma Rehabil 2004;19:205-16.
4. Dawson DR, Chipman M. The disablement experienced by traumatically brain-injured adults living in the community. Brain Injury 1995;9:339-53.
5. Ponsford JL, Olver JH, Curran C. A profile of outcome: 2 years after traumatic brain injury. Brain Injury 1995;9:1-10.
6. Lichtenberg PA, Stoltman J, Ficker LJ, Iris M, Mast B. A person-centered approach to financial capacity assessment: preliminary development of a new rating scale. Clin Gerontologist 2015;38:49-67.
7. British Columbia Ministry of Health/Public Guardian and Trustee of British Columbia. A guide to the certificate of incapability process under the Adult Guardianship Act. Victoria, British Columbia: British Columbia Ministry of Health; 2016. Available at: http://www.bit.ly/2tk7h4Y. Accessed June 15, 2015.
8. Walsh K, Bennett G. Financial abuse of older people. J Adult Prot 2000;2:21-9.
9. Dikmen SS, Machamer JE, Powell JM, Temkin NR. Outcome 3 to 5 years after moderate to severe traumatic brain injury. Arch Phys Med Rehabil 2003;84:1149-57.
10. Hartman-Maeir A, Soroker N, Ring H, Avni N, Katz N. Activities, participation and satisfaction one-year post stroke. Disabil Rehabil 2007;29:559-66.
11. Shigaki CL, Johnstone B, Schopp LH. Financial and vocational outcomes 2 years after traumatic brain injury. Disabil Rehabil 2009;31:484-9.
12. Koller K, Woods L, Engel L, Bottari C, Dawson D, Nalder E. Loss of financial management independence after brain injury: survivors' experiences. Am J Occup Ther 2016;70. 7003180070p1−8.
13. Webber L, Reeve R, Kershaw M, Charlton J. Assessing financial competence. Psychiatry Psychol Law 2002;9:248-56.
14. Engel L, Bar Y, Beaton D, Green R, Dawson D. Identifying instruments to quantify financial management skills in adults with acquired cognitive impairments. J Clin Exp Neuropsychol 2016;38:76-95.
15. Ghesquiere AR, McAfee C, Burnett J. Measures of financial capacity: a review. Gerontologist 2017:1-21.
16. Pinsker D, Pachana N, Wilson J, Tilse C, Byrne G. Financial capacity in older adults: a review of clinical assessment approaches and considerations. Clin Gerontol 2010;33:332-46.
17. Sousa LB, Simões MR, Firmino H, Peisah C. Financial and testamentary capacity evaluations: procedures and assessment instruments underneath a functional approach. Int Psychogeriatr 2014;26:217-28.
18. de Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide. Cambridge, UK: Cambridge University Press; 2011.
19. Mokkink LB, Terwee CB, Stratford PW, et al. Evaluation of the methodological quality of systematic reviews of health status measurement instruments. Qual Life Res 2009;18:313-33.
20. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;339:b2700.
21. Terwee C, Jansma E, Riphagen I, de Vet H. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res 2009;18:1115-23.
22. Feinstein AR. Clinimetrics. New Haven, CT: Yale University Press; 1987.
23. Marshall M, Lockwood A, Bradley C, Adams C, Joy C, Fenton M. Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia. Br J Psychiatry 2000;176:249-52.
24. Mokkink LB, Terwee CB, Patrick DL, et al. COSMIN checklist manual, version 9. Amsterdam, The Netherlands: COSMIN; 2012. Available at: www.cosmin.nl. Accessed May 5, 2015.
25. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34-42.
26. Terwee CB, de Vet HCW, Prinsen CA, Mokkink LB. Protocol for systematic review of measurement properties. Amsterdam, The Netherlands: COSMIN; 2011. Available at: http://bit.ly/2rZ25hJ. Accessed May 5, 2015.
27. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012;21:651-7.
28. Portney L, Watkins M. Foundations of clinical research: applications to practice. 2nd ed. Upper Saddle River, NJ: Prentice Hall Health; 2000.
29. Beaton DE, Bombardier C, Katz JN, Wright JG. A taxonomy for responsiveness. J Clinl Epidemiol 2001;54:1204-17.
30. Beaton DE. Understanding the relevance of measured change through studies of responsiveness. Spine 2000;25:3192-9.
31. van Tudler MW, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. Spine 2003;28:1290-9.
32. Rutman D, Silberfeld M. A preliminary report on the discrepancy between clinical and test evaluations of competence. Can J Psychiatry 1992;37:634-9.
33. Zur BM, Rudman DL, Johnson AM, Roy EA, Wells JL. Examining the construct validity of the Cognitive Competency Test for occupational therapy practice. Can J Occup Ther 2013;80:171-80.
34. Heaton RK, Marcotte TD, Mindt MR, et al. The impact of HIV-associated neuropsychological impairment on everyday functioning. J Int Neuropsychol Soc 2004;10:317-31.
35. Sadek JR, Stricker N, Adair JC, Haaland KY. Performance-based everyday functioning after stroke: relationship with IADL questionnaire and neurocognitive performance. J Int Neuropsychol Soc 2011;17:832-40.
36. Harvey PD, Raykov T, Twamley EW, Vella L, Heaton RK, Patterson TL. Factor structure of neurocognition and functional capacity in schizophrenia: a multidimensional examination of temporal stability. J Int Neuropsychol Soc 2013;19:656-63.
37. Marson D, Sawrie S, Snyder S, et al. Assessing financial capacity in patients with Alzheimer disease. Arch Neurol 2000;57:877-84.
38. Marson D. Loss of financial competency in dementia: conceptual and empirical approaches. Aging Neuropsychol Cogn 2001;8:164-81.
39. Triebel KL, Martin R, Griffith HR, et al. Declining financial capacity in mild cognitive impairment: a 1-year longitudinal study. Neurology 2009;73:928-34.
40. Kershaw M, Webber L. Assessment of financial competence. Psychiatr Psychol Law 2008;15:40-55.
41. Pachana NA, Byrne GJ, Wilson J, et al. Predictors of financial capacity performance in older adults using the Financial Competence Assessment Inventory. Int Psychogeriatr 2014;26:921-7.
42. Revheim N, Medalia A. The independent living scales as a measure of functional outcome for schizophrenia. Psychiatr Serv 2004;55:1052-4.
43. Revheim N, Schechter I, Kim D, et al. Neurocognitive and symptom correlates of daily problem-solving skills in schizophrenia. Schizophren Res 2006;83:237-45.

44. Baird A. (2006). Fine tuning recommendations for older adults with memory complaints: using the Independent Living Scales with the Dementia Rating Scale. Clin Neuropsychol 2006;20:649-61.
45. Green MF, Schooler NR, Kern RS, et al. Evaluation of functionally meaningful measures for clinical trials of cognition enhancement in schizophrenia. Am J Psychiatry 2011;168:400-7.
46. Velligan DI, Rubin M, Fredrick MM, et al. The cultural adaptability of intermediate measures of functional outcome in schizophrenia. Schizoph Bull 2012;38:630-41.
47. Quickel EJ, Demakis GJ. The Independent Living Scales in civil competency evaluations: initial findings and prediction of competency adjudication. Law Hum Behav 2013;34:155-62.
48. Brown C, Moore WP, Hemman D, Yunek A. Influence of instrumental activities of daily living assessment method on judgments of independence. Am J Occup Ther 1996;50:202-6.
49. Pickens S, Naik AD, Burnett J, Kelly PA, Gleason M, Dyer CB. The utility of the Kohlman evaluation of living skills test is associated with substantiated cases of elder self-neglect. J Am Acad Nurse Pract 2007;19:137-42.
50. Burnett J, Dyer CB, Naik AD. Convergent validation of the Kohlman Evaluation of Living Skills as a screening tool of older adults' ability to live safely and independently in the community. Arch Phys Med Rehabil 2009;90:1948-52.
51. Cramer K, Tuokko HA, Mateer CA, Hultsch DF. Measuring awareness of financial skills: reliability and validity of a new measure. Aging Ment Health 2004;8:161-71.
52. Marson D, Martin RC, Wadley V, et al. Clinical interview assessment of financial capacity in older adults with mild cognitive impairment and Alzheimer's disease. J Am Geriatr Soc 2009;57:806-14.
53. Heaton R, Rivera-Mindt M. Functional Impact Assessment training manual, version 1.1.
54. Knight A, Marson D. The emerging neuroscience of financial capacity. J Am Soc Aging 2012;36:46-52.
55. Lachs MS, Han SD. Age-associated financial vulnerability: an emerging public health issue. Ann Intern Med 2015;163:877-8.
56. Widera E, Steenpass V, Marson D, Sudore R. Finances in the older patient with cognitive impairment: "he didn't want me to take over." JAMA 2011;305:698-706.
57. Douglas A, Liu L, Warren S, Hopper T. Cognitive assessments for older adults: Which ones are used by Canadian therapists and why. Can J Occup Ther 2007;74:370-81.
58. Kramer JM, Schwartz A. Reducing barriers to Patient Reported Outcome Measures for people with cognitive impairments. Arch Phys Med Rehabil 2017;98:1705-15.
59. Lillie R, Kowalski K, Patry B, Sira C, Tuokko H, Mateer C. Everyday impact of traumatic brain injury. In: Marcotte T, I, Grant I, editors. Neuropsychology of everyday living. New York: Guildford Press; 2010. p 302-30.
60. Law MC, MacDermid J. Evidence-based rehabilitation: a guide to practice. 3rd ed. Thorofare, NJ: Slack Inc; 2014.
61. Loeb P. Independent Living Scales. San Antonio, TX: Psychological Corp; 1996.
62. Beaton DE, Bombardier C, Katz JN, et al. Looking for important change/differences in studies of responsiveness. J Rheumatol 2001; 28:400-5.
63. Moye J, Marson D, Edelstein B. Assessment of capacity in an aging society. Am Psychol 2013;68:158-71.
64. Gerstenecker A, Triebel K, Eakin A, Martin R, Marson D. Exploring the factor structure of financial capacity in cognitively normal and impaired older adults. Clin Gerontol 2018;41:33-41.
65. Kennedy CA, Beaton DE, Smith P. Measurement properties of the QuickDASH (Disabilities of the Arm, Shoulder and Hand) outcome measure and cross-cultural adaptations of the QuickDASH: a systematic review. Qual Life Res 2013;22:2509-47.
66. Mokkink LB, de Vet HC, Prinsen CA, et al. COSMIN Risk of Bias checklist for systematic review of patient-reported outcome measures. Qual Life Res https://doi.org/10.1007/s11136-017-1765-4 [E-pub ahead of print].
67. Dobson F, Hinman RS, Hall M, Terwee CB, Roos EM, Bennell KL. Measurement properties of performance-based measures to assess physical function in hip and knee osteoarthritis: a systematic review. Osteoarthritis Cartilage 2012:1548-62.
68. Francis DO, McPheeters ML, Noud M, Penson DF, Feurer ID. Checklist to operationalize measurement characteristics of patient-reported outcome measures. Syst Rev 2016;5:129.
69. Kottner J, Audige L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. Int J Nurs Stud 2011;48:661-71.
70. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2—assessing respondent understanding. Value Health 2011;14: 978-88.
71. Baird A, Podell K, Lovell M, McGinty S. Complex real-world functioning and neuropsychological test performance in older adults. Clin Neuropsychol 2001;15:369-79.
72. Bell-McGinty S, Podell K, Franzen M, Baird AD, Williams MJ. Standard measures of executive function in predicting instrumental activities of daily living in older adults. Int J Geriatr Psychiatry 2002; 17:828-34.
73. Christensen K, Mateer C, Williams R, Woodward T. Neuropsychological deficits, syndromes, and cognitive competency in schizophrenia. Cogn Neuropsychiatry 2005;10:361-78.
74. Dreer LE, DeVivo MJ, Novack TA, Marson DC. Financial capacity following traumatic brain injury: a six-month longitudinal study. Rehabil Psychol 2012;57:5-12.
75. Earnst KS, Wadley C, Aldridge T, et al. Loss of financial capacity in Alzheimer's disease: the role of working memory. Aging Neuropsychol Cogn 2001;8:109-19.
76. Gerstenecker A, Eakin A, Triebel K, et al. Age and education corrected older adult normative data for a short form version of the Financial Capacity Instrument. Psycholog Assess 2016;28: 737-49.
77. Griffith HR, Belue K, Sicola A, et al. Impaired financial abilities in mild cognitive impairment: a direct assessment approach. Neurology 2003;60:449-57.
78. Griffith HR, Stewart CC, Stoeckel LE, et al. Magnetic resonance imaging volume of the angular gyri predicts financial skill deficits in people with amnestic mild cognitive impairment. J Am Geriatr Soc 2010;58:265-74.
79. Kershaw M, Webber L. Dimensions of financial competence. Psychiatr Psychol Law 2004;11:338-49.
80. Kershaw M, Webber L. Financial Competence Assessment Inventory: Administration and Scoring Manual. Melbourne, Victoria, Australia: Matek Proprietary Limited; 2008.
81. Kohlman-Thomson L. The Kohlman evaluation of living skills. 3rd ed. Bethesda, MD: American Occupational Therapy Association; 1992.
82. Martin R, Griffith HR, Belue K, et al. Declining financial capacity in patients with mild Alzheimer disease: a one-year longitudinal study. Am J Geriatr Psychiatry 2008;16:209-19.
83. Martin RC, Triebel K, Dreer LE, Novack TA, Turner C, Marson D. Neurocognitive predictors of financial capacity in traumatic brain injury. J Head Trauma Rehabil 2012;27:E81-90.
84. Martin RC, Triebel KL, Kennedy RE, et al. Impaired financial abilities in Parkinson's disease patients with mild cognitive impairments and dementia. Parkinsonism Relat Disord 2013;19: 986-90.
85. Medalia A, Thysen J, Freilich B. Do people with schizophrenia who have objective cognitive impairment identify cognitive deficits on a self report measure? Schizophr Res 2008;105:156-64.
86. Mills WL, Regev T, Kunik ME, et al. Making and Executing Decisions for Safe and Independent Living (MED-SAIL): development and validation of a brief screening tool. Am J Geriatr Psychiatry 2014;22:285-93.

87. Okonkwo OC, Griffith HR, Vance DE, Marson DC, Ball KK, Wadley VG. Awareness of functional difficulties in mild cognitive impairment: a multidomain assessment approach. J Am Geriatr Soc 2009;57:978-84.
88. Okonkwo OC, Wadley VG, Griffith HR, Ball K, Marson DC. Cognitive correlates of financial abilities in mild cognitive impairment. J Am Geriatr Soc 2006;54:1745-50.
89. Revheim N, Medalia A. Verbal memory, problem-solving skills and community status in schizophrenia. Schizophr Res 2004;68:149-58.
90. Sherod MG, Griffith HR, Copeland J, et al. Neurocognitive predictors of financial capacity across the dementia spectrum: normal aging, mild cognitive impairment, and Alzheimer's disease. J Int Neuropsychol Soc 2009;15:258-67.
91. Stoeckel LE, Stewart CC, Griffith HR, et al. MRI volume of the medial frontal cortex predicts financial capacity in patients with mild Alzheimer's disease. Brain Imaging Behav 2013;7:282-92.
92. Triebel KL, Okonkwo OC, Martin R, Griffith HR, Crowther M, Marson DC. Financial capacity of older African Americans with amnestic mild cognitive impairment. Alzheimer Dis Assoc Disord 2010;24:365-71.
93. Van Wielingen L, Cramer K, Tuokko H. The Measure of Awareness of Financial Skills (MAFS): administration and scoring manual. Victoria, British Columbia, Canada: University of Victoria; 2004.
94. Van Wielingen LE, Tuokko HA, Cramer K, Mateer CA, Hultsch DF. Awareness of financial skills in dementia. Aging Ment Health 2004;8:374-80.
95. Wadley VG, Harrell LE, Marson DC. Self- and informant report of financial abilities in patients with Alzheimer's disease: reliable and valid? J Am Geriatr Soc 2003;51:1621-6.
96. Wang PL, Ennis KE, Copland SL. Cognitive Competency Test: manual. Toronto, Ontario, Canada: Mount Sinai Hospital Assessment & Rehabilitation; 1992.
97. Weiner MF, Gehrmann HR, Hynan LS, Saine KC, Cullum CM. Comparison of the test of everyday functional abilities with a direct measure of daily function. Dement Geriatr Cogn Dis 2006;22:83-6.
98. Zur B. Beyond the test manual of the Cognitive Competency Test (CCT). Occupational Therapy Now 2007;9:17-9.