

'Evaluation of testamentary capacity: A systematic review'

Himaja Aravind^a, Mark Taylor^{b,*}, Neeraj Gill^{a,b,c}

^a Mental Health and Specialist Services, Gold Coast Health, Gold Coast, QLD, Australia

^b School of Medicine and Dentistry, Griffith University, Gold Coast, QLD, Australia

^c Mental Health Policy Unit, Health Research Institute, University of Canberra, Canberra, ACT, Australia

ARTICLE INFO

Keywords:

Testamentary capacity
Testamentary ability
Assessment of capacity
Will
Financial decision making
Undue influence

ABSTRACT

Objective: To systematically review the literature on methods for the standardized and objective assessment of Testamentary Capacity (TC), to identify the best evidence-based and clinically pragmatic method to assess TC. Doubts concerning TC can have far-reaching legal and financial implications.

Method: A systematic search of the literature was conducted, using PRISMA guidelines, to identify studies which describe methods or tools for the assessment of TC.

Results: The Testamentary Definition Scale (TDS); the Testamentary Capacity Assessment Tool (TCAT); and the Testamentary Capacity Instrument (TCI) all have good psychometric properties, but TDS only partially assesses TC, and the TCI is designed for research rather than day-to-day clinical practice.

Conclusion: The TCAT could usefully supplement the clinical assessment of TC, coupled with a standardized examination of cognition. There is room to develop an all-encompassing TC assessment tool. Currently, the clinical judgement of a medical professional, taking account of the medical, legal, ethical issues informing a capacity or competency decision, remains the gold standard for assessing TC.

1. Introduction

The right of testation is a person's fundamental right, and Testamentary Capacity (TC), which is the degree of understanding the law requires a will-maker to possess, is a legal construct in most countries. A 'testator' (adult or older adult) is presumed by the law to be competent to make a will unless proven otherwise. The well-known *Banks vs. Goodfellow* case outlines the task-specific criteria for testamentary capacity (*Banks v Goodfellow, 1870*). The four main elements are a) the Testator's knowledge of the meaning and purpose of a will, b) the Testator's knowledge of the nature and extent of his/her property, c) the Testator's knowledge of his/her heirs and d) Testator's knowledge of the basic plan for the disposition of assets to heirs.

More recently, articles have been published attempting to update *Banks vs. Goodfellow* criteria to the current clinical context (*Bennett, 2016; Shulman, Himel, Hull, et al., 2017*). The elements primarily include whether the Testator is:

- Capable of understanding the act of making a will and its effects.
- Capable of understanding the nature and the extent of their property relevant to the disposition.

- Capable of evaluating the claims of beneficiaries, able to demonstrate an appreciation of the nature of any major conflict or complexity in the context of the Testator's life situation.
- Capable of communicating clear and consistent reasons for the distribution of their property, especially if there has been a relevant change from previously expressed wishes or prior wills; and.
- Free of a mental disorder, including delusions, that could influence the dispositive provisions of a will.

As highlighted by the case of *Spiers v Diane Hewston [2023] EWHC 1145*, the current thinking amongst lawyers on the continued relevance of the *Banks v Goodfellow* case varies according to jurisdiction - there may be no shared view of its relevance today although UK based lawyers still use *Banks v Goodfellow* as their starting point (C Pellow, personal communication). However, it is usually assumed by the Courts that a person has TC when he/she has sufficient cognitive capacity to understand the concept of the will, knowledge of his/her assets, awareness of who might have a claim on those assets, and ability to communicate the disposition of the estate after his/her death.

The United Nations Convention on the Rights of Persons with Disabilities (CRPD – see report from Essex Chambers), which includes long

* Corresponding author.

E-mail address: andrewmarktaylor55@gmail.com (M. Taylor).

term mental/intellectual impairments, affirms that persons with disabilities enjoy legal capacity on an equal basis with others in all aspects of life (article 12). The CRPD Committee emphasised the need to shift from substitute decision-making to supported decision-making, claiming that substitute decision-making regimes conflate mental capacity and legal capacity. They called for abolition of all substitute decision-making regimes including guardianship and mental health laws which allow for forced treatment, based on assessment of mental capacity. This interpretation of the CRPD has been termed as unrealistic or absolutist by many commentators, who have argued that legal capacity should be presumed, unless there is evidence to the contrary. Most countries have retained legal mechanisms like guardianship, mental capacity act or mental health acts which allow for substitute decision-making. However, the CRPD requires mechanisms to provide supported decision-making, so that substitute decision-making is a last resort.

Medical professionals, especially forensic psychiatrists, are commonly asked to provide an expert opinion about one's TC prospectively and retrospectively. This is expected to increase exponentially with an increasingly aging population, improving life expectancy, and a consequent increase in the prevalence of dementia, per capita wealth, and more complex family structures with increasing divorce and remarriage rates (Amanda, Ellen, Grace, et al., 2021; Kennedy, 2012; Purser, Magner, & Madison, 2015; Shulman, Cohen, & Hull, 2005).

The assessment of TC can be conducted by medical professionals (or lawyers) in various scenarios. For example, when a will is challenged by dispossessed heirs, when the legal personnel suspect incapacity, when there are allegations of incapacity due to cognitive impairment, mental illness, or substance use, or if there are concerns about the presence of undue influence, or in cases of posthumous assessments. Assessment of TC can be varied and challenging (Shulman et al., 2005) and are largely informal; there is no consistent approach. Hence the outcome can be subjective and inconsistent (Purser et al., 2015). While many standardized tools are available and regularly utilized in clinical settings to assess one's general capacity or capacity to consent to medical treatments (Sturman, 2005), there is a significant gap in the availability of standardized and objective measuring tools for TC.

Assessment of cognition is an essential element of TC assessment. Completing a will requires memory and executive functioning; hence, cognitive assessment is essential to TC assessment (Purser & Lonie, 2019). The process of completing a will can get more complicated if the testator's assets are vast and complex, and when there are multiple potential beneficiaries, the testator's demonstrable cognitive function becomes even more important. However, cognitive impairment does not always automatically preclude TC (Sousa, Simões, Firmino, et al., 2014). Severe cognitive impairment almost always rules out capacity; however, capacity can still be retained in mild or even moderate cognitive impairment. Historically Mini-Mental State Examination (MMSE), Montreal Cognitive Assessment (MoCA), Frontal Assessment Battery (FAB), and Addenbrooke's Cognitive Examination-Revised (ACE-R) are the most common cognitive assessment tools used as a part of capacity assessment. Each of these widely used instruments has drawbacks. For example, the MMSE lacks specificity as it does not measure long-term memory, memory recognition, or executive function (Molloy & Stanish, 1997) and it is also not sensitive to subtle cognitive impairments in psychiatric patients (Creavin, Wisniewski, Noel-Storr, et al., 2016; Faustman, Moses Jr., & Csernansky, 1990). MoCA is now gaining favour over the MMSE because of its greater sensitivity for Mild Cognitive Impairment (MCI) with less educational, language, and cultural bias (Ciesielska, Sokolowski, Mazur, et al., 2016; Lerner, 2012). None of these measures assess or predict TC directly.

Questions can be raised about TC in those with mental disorders, especially when there is a risk of its impact on cognition, for example, due to the recognized cognitive deficits in schizophrenia (Millan, Agid, Brüne, et al., 2012), which can impair capacity. In older patients with schizophrenia, cognitive decline resulting in dementia may be possible (Bergman-Levy, Heinik, & Melamed, 2014). However, the mere

diagnosis of dementia (any type) or mental illnesses may not be sufficient to deprive a patient of his or her TC. Dementias and mental illness can be significant risk factors for incapacity but do not inevitably lead to incapacity. Therefore, completing a mental state examination as part of the TC assessment is vital.

Wills can be contested if concerns about undue influence (coercion) exist. The risk is increased in the elderly due to the risk of cognitive impairment. Various commentators discuss the relevance of considering and ruling out undue influence as a part of TC assessment (Marson, Huthwaite, & Herbert, 2004; Stankowski & Noffsinger, 2005; Kenneth I. Shulman, Carole, Felice, Kirsh, et al., 2007; Peisah, Finkel, Shulman, et al., 2009).

Although not directly related, impaired financial knowledge and skills can raise questions about TC. Assessing the testator's knowledge of their financial assets and assessing financial capacity can be an essential aspect of TC assessment (Sousa et al., 2014; Wood & Lichtenberg, 2017).

Doubts concerning TC can have far-reaching legal, financial, and interpersonal implications. Thus, a rigorous understanding of the causes of incapacity and a reliable, valid, and evidence-based measure of capacity becomes vital, and identifying objective and quantitative tools will assist clinicians and the legal system when wills or one's TC is contested.

1.1. Objective

This study aims to systematically review methods for the standardized and objective assessment of TC, and identify the best evidence-based and clinically pragmatic method to assess TC.

2. Methodology

2.1. Information sources and search strategies

A systematic search of the literature was conducted according to PRISMA guidelines to identify studies in which methods or tools are described that may be used for the assessment of one's TC.

The databases used included CKN, Cochrane, PubMed, CINAHL, MEDLINE, Embase, APA PsycINFO, and Google Scholar. Keywords were used to ensure the inclusion of all studies that used suitable methods or tools. Key words included 'testamentary capacity, testamentary ability, assessment of capacity, will, financial decision making, undue influence.'

2.2. Selection process

- **Inclusion criteria:** All studies published in English, irrespective of the year published, were included if they described a TC assessment method or an assessment tool of TC in detail (component measured, type of measure, length of the assessment and psychometric properties)
- **Exclusion criteria:** Articles that did not describe a TC assessment method; articles that did not contain original data; and review and opinion articles were excluded.
- **Selection of relevant articles:** The lead author reviewed the databases. Duplicates, irrelevant articles, titles without abstracts, and non-English articles were omitted. Two authors (HA and MT) independently examined the title and abstract of the remaining articles, and any articles irrelevant to the study were removed. For those studies in which inclusion criteria were met, the entire article was sourced and assessed for relevancy.
- **Data extraction and synthesis:** The included articles were critically appraised independently by authors (HA and MT) to decide on the studies to be included in the final analysis. Relevant data were synthesized, and results were incorporated into tables where appropriate. A meta-analysis was impossible due to the nature of the

‘intervention’ (diagnostic tool) and ‘outcome measure’ (TC assessment).

3. Results

3.1. Search results

The initial search of the databases produced 283 references. After initial screening, where duplicates, irrelevant articles, titles without any abstracts, and non-English articles were omitted, 60 articles remained. After careful review (by HA and MT independently) of the 60 abstracts only 11 studies met inclusion criteria. The full text of these 11 included studies was sourced and critically appraised. Of the 11 included studies, one did not introduce a method or a tool in sufficient detail (Shulman, Peisah, Jacoby, et al., 2009). Another introduced a tool that primarily focused on financial awareness and susceptibility and not TC (Lichtenberg, Gross, & Campbell, 2020). Data were then extracted from the remaining nine studies. This process is illustrated in Fig. 1 and the 11 included studies are summarised in Table 1.

3.2. Overview of the included studies

3.2.1. Test for creative abstraction

Eliasberg (1953) was the first to publish an article proposing a standard method to check TC by assessing for ‘soundness/unsoundness of mind,’ and by ruling out undue influence. Eliasberg highlighted the

importance of gathering collateral information, assessing the handwriting of the testator, and ruling out neurological conditions which could influence mental capacity and the ability for testament temporary or permanently. A ‘test for creative abstraction’, was introduced based on the mental function of ‘abstraction,’ using slips of paper of assorted colours with a cigarette on the reverse side of one slip. It tries to establish the ability of aphasics (including young children who have not yet developed full language ability) to interpret different scenarios, demonstrating the preservation of general intelligence and orientation despite the loss of language ability. This test can be seen as an attempt to establish TC in brain-injured or aphasic patients.

This study has significant limitations. There is no clear description of the study group or the description of the actual test itself. The scoring of the test is not clear. The psychometric properties of the test were not explored or established.

3.2.2. Testamentary Definition Scale (TDS)

Heinik, Werner, and Lin (1999) introduce a test that evaluates and quantifies the ability to define “testament,” one specific component of TC. The test asks the person to define testament, and the answer is documented as verbatim. The answer is scored against six identified items (1) a testament as a document, (2) a person makes a testament, (3) a testament is made during a person’s lifetime, (4) a testament involves property, (5) a testament involves a receiver or receivers [heirs], and (6) a testament comes to fruition after death. In the study, the most common elements of a “testament” identified by participants included that it

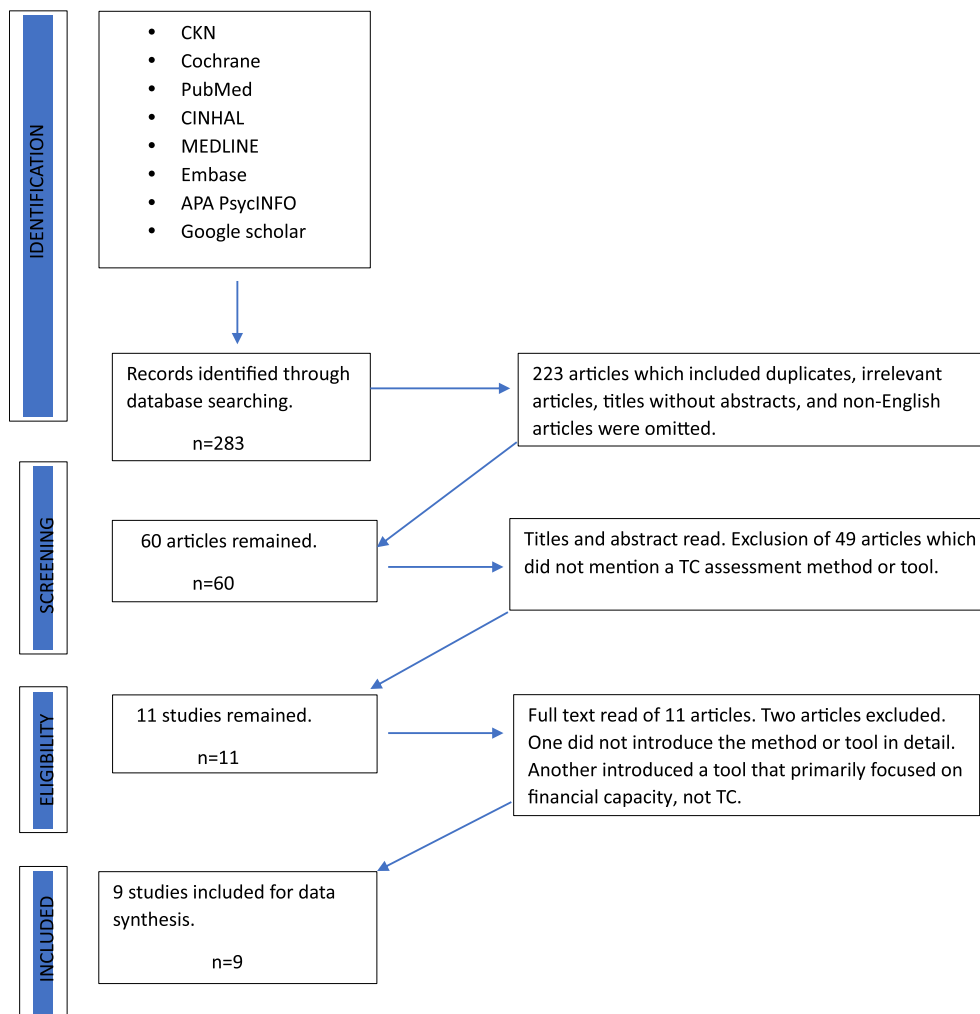


Fig. 1. Search results summary Flow chart.

Table 1
Summary of the studies.

Year & place of study	Sample	Instruments or methods examined	Results	Limitations	Comments
1953, USA	Unknown (Says large group of children of different ages, adolescents, adults, and older adults with and without dementia).	Test for creative abstraction.	Establish TC in brain-injured/aphasic patients. Demonstrates preservation of general intelligence and orientation despite loss of language ability. Tries to establish that intelligent actions are independent of linguistics.	No clear description of the study group. No case/ control specified. No clear description of the actual test itself. Scoring the test not clear. Psychometric properties not explored.	Not a clinically useful tool to assess TC.
1999, ISRAEL	31 community dwelling elderly individuals (61–85 years) referred to a psychogeriatric OP (Outpatient).	Testamentary Definition Scale (TDS).	Evaluates and quantifies the ability to define testament, one specific component of TC in elderly (>60) with cognitive impairment. Reliability (interrater reliability 0.77, internal consistency 0.76) and validity (consensus-100%, criterion- mean score of 3.9 and 1.6 for an MMSE score of 24–30 and 0–23 respectively) established. Cut off scores established. 8% FP (False Positive), 94% PPV (Positive Predictive Value). Statistically significant correlations found for MMSE & CAMCOG (Cambridge Cognitive Examination) but not when controlled for education, age, and sex.	6 items were based on an encyclopedia, not a widely accepted legal definition (e.g. Banks vs Goodfellow) Small homogenous sample, 24/31 had dementia. No control groups. No comparison with healthy older adults. Possible incidental consensus validity, based on judgement of five old age psychiatrists. Correlation with other cognitive domains and legal requirements not explored.	Partial assessment tool but potential clinical utility.
2008, UK	74 patients (>55 yrs) from old age psychiatric team with a diagnosis of AD (Alzheimer's Dementia).	Assessment of TC based on guidelines outlined by the BGS (British Geriatric Society), BMA (British Medical Association) and The Law Society	TC was mostly established in mild (62.5%) and moderate (35%) AD. 2.5% of severe AD had TC. Using logistic regression, two global cognitive screening measures [the MMSE and CAMDEX-R (Cambridge Examination for Mental Disorders of the Elderly)] equivalently predicted participants' TC status in 87% of the participants. Language measures predicted TC status in 84% of the participants, whereas impaired short-term memory and poor concentration were not accurate predictors on TC. The structured interview had high interrater reliability (0.82).	An "independent investigator" was trained to assess TC. Interviews were then independently assessed by a second rater, an old age psychiatrist. Unclear what interview information the first investigator provided to the second rater. ~2/3rd of the sample was female. No control groups. The structured interview and its focus domains were not described.	The study recommends assessment of TC based on legal criteria in cognitively healthy individuals and an assessment of TC along with MMSE for cognitively impaired individuals. This can be adopted in a clinical setting.
2008, ITALY	25 consecutive individuals referred to neuropsychology with suspected cognitive impairment.	Handwriting scale	Explores correlation of handwriting with cognitive deterioration. Semiquantitative score system→verbal, lexical and spatial orientation. Very good inter-rater agreement. Significant correlation identified between writing score, MODA (Montreal Dementia Assessment) and MMSE for all parameters analysed. No patient who had a writing score ≤ 5 had a MODA score higher than 60 or a MMSE higher than 20. Thus, a writing score ≤ 5 is a strong indicator of severe mental impairment.	Sociodemographic factors of the sample not described. MODA is not used as a standard cognitive assessment apart from in Italy. Study did not explore the impact of education on writing score.	Potentially useful tool for retrospective assessment. Relevant in countries where the law requires the will to be handwritten. Not suitable for patients with severe motor deficit in the dominant hand. May not be automatically applied to people that had known focal brain lesions.
2018, CANADA		Contemporaneous Assessment Instrument (CAI)	Uses MacCAT-T (MacArthur Competence Assessment Tool for Treatment), as a gold standard for CAI, and recommends incorporation of legal criteria (a semi-structured interview) + neuropsychological assessment including working memory and language, as	A systematic search of electronic data base and review of relevant articles and books from reference lists, as well as articles that highlighted instruments for assessing other decisional capacities. No study group or case/ control.	Potentially useful template in clinical setting for contemporaneous assessment of TC rather than an assessment tool.

(continued on next page)

Table 1 (continued)

Year & place of study	Sample	Instruments or methods examined	Results	Limitations	Comments
2018, Greece	64 individuals with mild to moderate dementia, from cognitive disorder clinic.	Testamentary Capacity Assessment Tool (TCAT)	well as specific executive functions such as judgement, planning, and reasoning. Also recommends the involvement of the testator and his or her collateral, review of evidence (medical records or related correspondences not specified in previous wills), legal documents, and witness testimony. Assess TC in people with Mild to moderate Dementia (Alzheimer's type predominantly). Assesses memory (including autobiographical), absence of serious psychopathology, knowledge of financial parameters and intention. Taps cognitive domains more specific for TC. Forensic psychiatrist, whose assessment was the gold standard, was blind to TCAT results. Psychometric properties well studied. Cronbach's alpha showed high levels of internal consistency reliability (0.86), the correlation coefficients showed high levels of criterion-related validity (0.797, $p < 0.001$). Satisfactory sensitivity and specificity established for different cut off scores for different versions. Comparison with the MMSE showed a clear advantage of the TCAT both in terms of sensitivity and specificity, supporting a superior validity.	No test as such so no scoring. Psychometric properties not explored.	Potentially useful tool in clinical setting – brief, good statistical evidence for reliability and validity, direct assessment tool, assess the person's core functions which are required for TC.
2021, ITALY	323 (123 males and 200 females) healthy Italians of different ages (31–93) and educational background.	Testamentary Capacity Assessment Tool (TCAT)	TCAT validated in healthy males and females of a range of age and education. Provides normative data and cut off values, sociodemographic factors affecting scores. Regression analyses revealed a significant effect for gender, age and education on TCAT scores. Significant associations between the TCAT and the MMSE, MoCA, FAB and BDI-II (Beck Depression Inventory) A positive correlation between the TCAT and TDS was also found, providing good convergent validity of the TCAT and TC.	Under 40 poorly represented. More females than males in sample.	Potentially useful tool in clinical setting – brief, good statistical evidence for reliability and validity. Directly assesses the core functions required for TC.
2022, USA	20 (mild to moderate AD) +22 cognitively intact older adults >60.	Testamentary Capacity Instrument (TCI)	Assess TC, and differentiate cognitively intact elderly from cognitively impaired elderly. Assesses cognition and mood as well. TCI is administered verbally or in writing. An overall quantitative score emerges as capable, marginally capable, or incapable. 4 TC elements were selected based on the Anglo-American legal literature, derived from Banks v. Goodfellow. Good internal consistency for two elements and good inter-rater reliability. Good face, content, and construct validity.	Small sample size. Four core components of TC likely do not represent the sole conceptual sources or frames of reference about TC under Anglo-American law. Tool developed primarily for forensic experts and not all medical professionals.	Potentially useful tool in clinical setting for forensic experts.
2022, USA	20 (mild to moderate AD) +22 cognitively intact older adults >60.	Testamentary Capacity Instrument (TCI)	Explored cognitive correlates of TC by applying TCI in controls and AD. The act of testation is highly affected by cognitive ability associated with measures of general cognition, language, memory, executive function, and knowledge speed.	Small sample size. Four core components of TC do not represent the sole conceptual sources regarding TC under Anglo-American law.	Potentially useful tool in clinical setting for forensic experts.

involved a person (58%) and that it took effect after a person's death (45%). <40% of the sample identified the other key four criteria of testament. Psychometric properties were explored. Reliability (inter-rater reliability 0.77, internal consistency 0.76) and validity (consensus-100%, criterion- mean score of 3.9 and 1.6 for an MMSE score of 24–30 and 0–23, respectively) were established. Calibrated and cut-off scores established. At a score of 0–2, 8% False Positives, 94% Positive Predictive Value. Statistically significant correlations were initially found for MMSE and CAMCOG (Cambridge Cognitive Examination) but disappeared when controlled for education, age, and sex.

The study has limitations. The six items were based on the Encyclopedia definition of TC but not a widely accepted legal definition (e.g., Banks vs. Goodfellow). The sample was small and homogenous; the influence of age or education on the ability to define testament was not tested. Correlation with other cognitive domains and legal requirements was also not explored. A consensus validity of 100%, which was based on a judgement of five old age psychiatrists, could be incidental or biased. There was no study on a control group without dementia.

3.2.3. Assessment of TC based on guidelines outlined by the British Geriatric Society (BGS), British Medical Association (BMA) and The Law Society

Roked and Patel (2008) discuss a method of assessing TC in people with dementia based on the guidelines provided by British Geriatric Society, British Medical Association and The Law Society, which in turn is based on the Banks vs Goodfellow criteria. An “independent investigator” was trained to assess TC using the guidelines via structured interviews of participants, and these interviews were then independently assessed by a second rater, an old age psychiatrist. The results showed that TC was primarily established in mild AD where 6% of mild AD patients, 38% of moderate AD patients, and 56% of severe AD patients lacked TC. Using logistic regression, two global cognitive screening measures - the MMSE and CAMDEX-R (Cambridge Examination for Mental Disorders of the Elderly) equivalently predicted participants' TC status in 87% of the participants. Amongst the cognitive domain functions, language measures predicted TC status in 84% of the participants, whereas impaired short-term memory and poor concentration were not accurate predictors of TC. The study recommends an assessment of TC based on legal criteria in cognitively healthy individuals and an assessment of TC along with MMSE for cognitively impaired individuals. The structured interview had high interrater reliability (0.82).

It was observed that two thirds of the sample were female. It is unclear what interview information by the first investigator was provided to the second rater. The structured interview and its focus domains were not described.

3.2.4. Handwriting scale

Fontana, Dagnino, Cocito, et al. (2008) introduced a test for retrospective assessment of TC. It explores the correlation of handwriting with mental deterioration. The test is a semiquantitative scoring system to assess verbal, lexical, and spatial orientation. Participants were asked to write a spontaneous text of their choice, about six to seven lines in length, written on a blank sheet; a text dictated by the examiner that the patient was requested to write on a boxed sheet; the same text, dictated by the examiner, that the patient was requested to write on a blank sheet. Verbal and lexical skills together were scored from 0 to 5. Spatial orientation alone was scored between 0 and 5. Three authors independently scored each writing. The sum of both was calculated as the writing score. Psychometric properties showed very good inter-rater agreement (0.8 and 0.94). A significant correlation was identified between the writing score, MODA (Montreal Dementia Assessment), and the MMSE for all parameters analysed. No patient who had a writing score ≤ 5 had a MODA score higher than 60 or an MMSE higher than 20. Thus, a writing score ≤ 5 is a strong indicator of mental impairment.

There was no evidence available for the influence of education on writing scores. The tool may be relevant in countries where the law

requires that the will be handwritten but may not be a suitable assessment for individuals with severe motor deficits in the dominant hand. Additionally, MODA is only used as a standard cognitive assessment in Italy.

3.2.5. Contemporaneous Assessment Instrument (CAI)

Brenkel and his colleagues (Brenkel, Whaley, Herrmann, et al., 2018) introduced CAI as a method of TC assessment. It uses MacCAT-T (MacArthur Competence Assessment Tool for Treatment), a widely used instrument for assessing medical decision-making, as a gold standard for CAI. CAI was proposed after a systematic search of an electronic database (1946 to 2016), with a review of relevant articles and books from reference lists, and articles that highlighted instruments for assessing other decisional capacities.

CAI is a structured approach covering the range of abilities necessary for establishing TC. It uses a semi-structured interview with standardized criteria for the legal test along with a brief validated neuropsychological assessment including working memory and language, as well as specific executive functions such as judgement, planning, and reasoning. CAI involves the testator and his or her collateral, and a review of the evidence, legal documents, and witness testimony.

This study's limitations include no description of an evaluation of CAI in a normal or clinical population. The scoring of the method needs to be clarified. Psychometric properties were not explored.

3.2.6. Testamentary Capacity Assessment Tool (TCAT)

Papageorgiou et al developed TCAT to assess TC in people with mild to moderate Dementia of Alzheimer's type (Bolognini, Gramegna, Esposito, et al., 2021; Papageorgiou, Voskou, Economou, et al., 2018). They adopted a case-control model for their study. A forensic psychiatrist who was blind to the TCAT results and cognitive status was the gold standard. The tool assessed memory (including autobiographical), absence of serious psychopathology, knowledge of financial parameters, and intention - cognitive domains more specific for TC. Psychometric properties were well studied. Sensitivity and specificity were obtained for different cut-off scores. Cronbach's alpha showed high levels of internal consistency (0.86), and the correlation coefficients showed high levels of criterion validity (0.797, $p < 0.001$). Comparison with the MMSE showed a clear advantage of the TCAT both in terms of sensitivity and specificity, supporting the superior validity. The TCAT has satisfactory psychometric properties in individuals with low educational levels. TCAT was further validated in a larger group of healthy males and females of a range of ages and education (Bolognini et al., 2021). It provides normative data and cut-off values, sociodemographic factors affecting scores. Multiple regression revealed a significant effect of gender, age, and education on TCAT scores. Correlation analyses showed significant associations between the TCAT and the MMSE, MoCA, FAB, and BDI-II (Becks Depression Inventory). A positive correlation between the TCAT and TDS was also found, proving good convergent validity of the TCAT with respect to TC.

It is unclear whether the TCAT adopted an accepted theoretical model of TC assessment. The sample in the original study was predominantly females, and those aged under 40 were poorly represented (neurological disorders with juvenile-onset). A single expert served as the “gold standard.” No results were available regarding possible fluctuations in cognitive functions. Some of the items employed in the TCAT are specific to the Greek population, although they can easily be modified to other cultural settings.

3.2.7. Testamentary Capacity Instrument (TCI)

Marson et al introduce the TCI (Gerstenecker, Martin, Hebert, et al., 2022; Martin, Gerstenecker, Hebert, et al., 2022). The TCI aims to assess TC and differentiate cognitively intact older adults from cognitively impaired older adults. It was specifically designed to assess TC based on four legal elements: (a) understanding of the nature and purpose of a will (purpose), (b) recollection of the nature and extent of one's property

(assets); (c) memory for potential heirs or objects of one's bounty (heirs); and (d) rational plan for the disposition of assets (plan). Elements were selected based on a review of current Anglo-American legal requirements for TC across the 50 states in the United States and the senior author's forensic knowledge and experience. In addition, individual items were designed and reviewed in collaboration with psychologists and attorneys knowledgeable of TC and dementia populations. Items were then grouped according to the construct or legal element they believed to measure. The items of the TCI are administered orally or in writing, and an overall performance score emerges in one of three categories: capable, marginally capable, or incapable. It assessed older adults >60 yrs. of age and was a case-control study (subjects were either tested cognitively intact or had Alzheimers). The test demonstrated good psychometric properties with good internal consistency, good inter-rater reliability, as well as adequate face, content, and construct validity. The further validation study identified that testation is highly affected by general cognition, language, memory, executive function, and knowledge speed.

The sample size is small. The tool is designed for forensic experts and involves a long interview and the need for collateral history.

4. Discussion

4.1. Main findings

This systematic review aimed to identify the best evidence-based and clinically pragmatic method (a time efficient, easily accessed, user friendly, validated tool) to assess TC, whilst incorporating all elements (cognition, financial capacity, knowledge of testament, mental status, and presence of undue influence) relevant to TC.

The available literature was limited and of varying quality. It was noted that much of literature on assessing TC recommends using a simple cognitive assessment such as the MMSE or MoCA (or ACE-R if time permits) in conjunction with a clinical review of the individuals. However, six specific methods or tools for prospective assessment and one (handwriting scale) for retrospective assessment of TC were also identified. Only one included study (Eliasberg, 1953) attempts independence from written or linguistic skills, although impairment of these skills itself could lead to testamentary incapacity. All the included instruments are summarised in Table 2.

Three of the six identified tools for prospective assessment, TDS, TCAT, and TCI, demonstrated good statistical evidence, and are possibly viable candidate TC assessment tools. Assessment of TC based on guidelines outlined by the BGS, BMA and The Law Society had high inter-rater reliability however the 'structured interview' was not described adequately. CAI was not supported by any of its psychometric properties. All the studies except the test for creative abstraction rely directly on hand-writing and linguistic ability. It must be assumed that writing and linguistic ability plays a direct role in an individual's testamentary ability.

TDS developed in Israel and was tested in 31 people. It evaluates the cognitively impaired elderly individual's understanding of wills and will-making and scores it against six item scale. It is time efficient, user friendly, and easy to score. The tool is reliable, and valid. It also gives cut off scores. There is no need for collateral history. However, it did not specifically aim to examine TC itself and is based on an encyclopedic definition of TC and not on any accepted theoretical model. The tool is also not validated in healthy individuals. Demographic, cultural, and ethnic factors could influence the scores via the understanding of testament. It cannot be recommended as a stand-alone tool.

TCAT devised in Greece and tested in Italy, and has three assessment strands, namely autobiographical memory; financial knowledge and awareness; and a theory of mind section looking at intention and judgement. It assesses TC in people with mild to moderate dementia (Alzheimer's type predominantly). It is time efficient (15–20 min, as per the authors), easy to administer, requires no special training, and is

easily assessed by contacting the author or through the journal. It has a flexible scoring system according to the social situation, and there is no need for collateral information. It is reliable and validated in both cases and in healthy individuals. The tool taps cognitive domains more specific for TC and can be used by experts and non-experts. Although not scored within the tool itself, it rules out depression contributing to the results by administering a PHQ-9 (Patient Health Questionnaire). The tool, as developed in Greece, will need minor modification to be suitable for other countries. It is also not clear whether the tool is based on any accepted theoretical model (e.g Banks vs. Goodfellow).

The TCI from USA has four 'elements' based on Anglo-American legal literature for TC and Banks vs Goodfellow criteria. It focuses on cognitive domains relevant for each element. It is reliable and valid. It is intended for forensic experts and is validated in healthy individuals. Overall performance score emerges in one of three categories: capable, marginally capable, or incapable. However, it is a long interview instrument which needs collateral history as well as other neuropsychological assessment. The tool gives no cut off scores and does not confirm or rule out TC. Ultimately, it requires a clinicians' clinical judgement to assess TC.

Authors of TCI and TCAT were contacted directly to access the tool for better visual comparison however only TCAT was able to be accessed. The author of TCI has confirmed that the tool is developed for research purposes and is not available for clinical use currently.

Comparison between TDS, TCAT and TCI is demonstrated in Table 3.

4.2. Summary of the findings

Most of the identified tools for TC assessment rely on the theoretical concept of the case of *Banks Vs Goodfellow*, which is adopted in the legal system of many Anglo- American countries. For prospective assessment of TC, TCAT and TCI are more specific and demonstrate overall better performance, despite both having differing drawbacks and not meeting our full criteria for recommendation as an ideal tool which combines a valid evidence base with clinical pragmatism. However, TCI is a research tool and is not available for clinical use, which leaves only TCAT for day-to-day clinical practice.

4.2.1. Strengths and limitations of this review

Our study had a clear objective. It included all relevant articles published without imposing a time limit, thus identifying one of the earliest articles - published in 1953. This systematic review provides insight into the advancing quality of studies exploring TC tools and used a priori criteria to make a firm recommendation that is useful for lawyers and clinicians.

However, the number of included studies was small. The sample size in all the studies was small and homogeneous. Unpublished and non-English articles were not included in the review, and most of the included studies included were completed either in USA or Europe which creates a geographical bias. A gender imbalance in study publications (with women more often represented than men) was noticeable, as well as a frequent lack of description of the study population in the selected studies, which can also be considered as a limitation.

4.3. Recommendations

4.3.1. Implications for research

There are inherent difficulties in conducting research in this area (e.g., recruiting sufficient samples, obtaining consent, developing a tool with specific domains relevant to legal and medical professionals, and ethical concerns while administering an experimental method or tool in cognitively impaired elderly). Nevertheless, more empirical research is undoubtedly needed into this area to develop a time efficient, easily accessed, user friendly, validated tool incorporating all the significant elements which could influence one's testamentary capacity. Further studies with a large heterogeneous sample are recommended. Most of

Table 2
Measuring abilities of the instruments.

Instrument	Component measured	Type of measure	Length	Psychometric properties	Types of samples	Comments
Test for creative abstraction	Demonstrates preservation of general intelligence and orientation despite loss of language ability	unclear	Not stated	Not available	'Large group' of children of different age, adolescents, adults, and older adults with and without dementia n = unclear	Not a time efficient, easily accessed, validated tool. Does not incorporate all significant TC elements.
Testamentary Definition Scale (TDS)	Evaluates and quantify the ability to define testament, one specific component of TC.	Asks the patient to define 'testament' and recording it verbatim/self-report.	Not stated specifically however appears to be short as only 6 items.	High interrater reliability, internal consistency, and validity. Calibrated and cut off scores established.	Community dwelling elderly (>60 yrs) with cognitive impairment n = 31	Time efficient, easily accessed, friendly, validated tool. However, based on Encyclopedia definition and only checks ability to define testament.
Assessment of TC based on guidelines by the BGS, BMA and The Law Society	Contemporaneous assessment of TC.	Structured interview	Not stated.	High interrater reliability (k 0.82)	Sample > 55 yrs. with mild, moderate, and severe AD. ~2/3 rd. female n = 74	The structured interview and its focus domains were not described.
Handwriting Scale	Retrospective/posthumous assessment of TC. Correlation of handwriting with cognitive deterioration.	Semiquantitative standardized score system which assess verbal, lexical and spatial orientation	Not stated	Very good inter-rater agreement. Significant correlation between writing score, MODA and MMSE.	Those referred to Neuropsychology Service for clinical suspected 'mental deterioration'. n = 25	Not relevant for prospective assessment but possible tool for retrospective assessment of TC.
Contemporaneous Assessment Instrument (CAI)	Full range of abilities necessary for establishing TC.	A semi-structured interview + neuropsychological assessment	Not stated.	Not explored or explained.	Not stated	Possible use in a clinical setting. Not a tool for assessing TC.
Testamentary Capacity Assessment Tool (TCAT)	Contemporaneous/prospective assessment of TC. Assess the person's core functions which are required for TC.	Questionnaire	15–20 min	Cut off scores showed high reliability, sensitivity, specificity, criterion validity.	Individuals with mild to moderate Dementia (AD). n = 64	Time efficient, friendly, easily assessed, but not recommended as a solo clinical tool.
Testamentary Capacity Instrument (TCI)	Contemporaneous/prospective assessment of TC. Assess TC and differentiate cognitively intact elderly from cognitively impaired elderly.	Long interview-based test	Not stated in the study but appears lengthy	Good internal consistency, inter-rater reliability, good face, content, and construct validity	Older adults >60 n = 44	Mainly aimed for forensic experts. Not time efficient or easily accessed. Not a solo tool.

the identified tools rely on the theoretical concept of the case of Banks vs Goodfellow, which is adopted in the legal system of many Anglo-American countries, however their utility in other countries with significant social, economic, educational, and cultural differences is unknown. It is also important to consider that most of the existing studies were done either in USA or Europe, and awareness about the significance of TC could potentially vary between different countries and cultures, so more linguistically and culturally diverse studies are needed.

4.3.2. Implications for clinicians

The evidence available confirms that there is no standardized stand-alone tool, which is time efficient, easily accessed, friendly and validated incorporating all the elements (cognition, financial capacity, knowledge of testament, mental status, and presence of undue influence) that is highly relevant to TC. It is important for medical professionals to be aware of TCAT, and consider using it as a part of a comprehensive evaluation of TC.

Administration of these tools will inevitably assist medical professionals including forensic psychiatrists to gather valuable information regarding ones TC however a neurocognitive assessment (ideally focusing on executive function, verbal memory, and semantic knowledge) and a detailed mental state examination is recommended as an adjunct for a complete assessment of TC which also incorporates

collateral, a review of evidence, legal documents, and witness testimony.

5. Conclusion

TCAT would usefully supplement the process of assessing TC, along with a cognitive assessment like MMSE or MoCA, in a busy everyday clinical practice. More research is required to create a stand-alone tool incorporating all the significant elements that could influence one's TC. The clinical judgement of a medical professional, which considers the medical, legal, ethical, and other factors that inform a competency decision, remains the gold standard for TC assessment.

CRedit authorship contribution statement

Himaja Aravind: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mark Taylor:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Neeraj Gill:** Writing – review & editing, Writing – original draft, Supervision, Project administration.

Table 3
TDS vs. TCAT vs. TCI.

	TDS	TCAT	TCI
STRENGTHS	Assesses TC in people with cognitive impairment. Time efficient (6 items), user friendly, easy to score. Reliable, and valid. Inter-rater reliability. Calibrated and cut off scores established. No need for collateral history.	Assess TC in people with Mild to moderate Dementia (Alzheimer's type predominantly). Time efficient, friendly, easy to administer, no special training needed, easily assessed by contacting the author or through the journal. Direct assessment tool without the need for collateral history. Flexible scoring system according to social situation. Reliable and valid. The TCAT has satisfactory psychometric properties in individuals with low educational level, and was further validated in healthy men and women. Taps cognitive domains more specific for TC. Can be used by expert and non-expert. Rules out depression contributing to the results	Designed for forensic experts. 4 'elements' based on Anglo-American legal literature for TC and Banks vs. Goodfellow. Focus on cognitive domains relevant for each element. Good internal consistency for two elements (two other elements inherently varied across participants), inter-rater reliability, good face, content, and construct validity. Validated in healthy individuals. Overall performance score results in one of three categories: capable, marginally capable, or incapable.
LIMITATIONS	Not validated in healthy individuals. Checks only one aspect of TC (ability to define testament) in cognitively impaired individuals.	May need further clinical evaluation in complex scenarios (e.g.: complex mental health issues, complex assets). Not clear whether its based on a theoretical model (e.g Banks vs. Goodfellow).	Research tool and not available for clinical use. Long interview instrument. Needs collateral history. Will still need other neuropsychological assessment. No cut off scores and does not confirm or rule out TC. Needs clinicians' clinical judgement to assess TC.

Declaration of competing interest

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors received no financial support for the research, authorship, and/or publication of this article. Ethical approval was not required for this review.

Acknowledgement

The authors acknowledge the input from the librarians at Gold Coast University Hospital for resourcing materials needed for the study, and to Caroline Pellow and Rosie Taylor of Womble Bond Dickinson.

References

- Amanda, K., Ellen, J., Grace, J. L., et al. (2021). A comprehensive approach to assessment of testamentary capacity. *Frontiers in Psychology*, 12.
- Banks v Goodfellow. (1870). *LR 5 QB 549. Queen's bench division (QBD)*.
- Bennett, H. (2016). *Neurolaw and Banks v Goodfellow (1870): Guidance for the assessment of testamentary capacity today. Australasian Journal on Ageing*, 35(4), 289–290.
- Bergman-Levy, T., Heinik, J., & Melamed, Y. (2014). Testamentary capacity of the schizophrenic patient. *Israel Medical Association Journal*, 16(3), 142–146.
- Bolognini, N., Gramegna, C., Esposito, A., et al. (2021). The testamentary capacity assessment tool (TCAT): Validation and normative data. *Neurological Sciences*. <https://doi.org/10.1007/s10072-021-05736-8>
- Brenkel, M., Whaley, K., Herrmann, N., et al. (2018). A case for the standardized assessment of testamentary capacity. *Canadian Geriatrics Journal*, 21(1), 26–31.
- Ciesielska, N., Sokolowski, R., Mazur, E., et al. (2016). Is the Montreal cognitive assessment (MoCA) test better suited than the Mini-mental state examination (MMSE) in mild cognitive impairment (MCI) detection among people aged over 60? Meta-analysis. *Psychiatr Pol*, 50(5), 1039–1052.
- Creavin, S. T., Wisniewski, S., Noel-Storr, A. H., et al. (2016). Mini-mental state examination (MMSE) for the detection of dementia in clinically unevaluated people aged 65 and over in community and primary care populations. *COCHRANE DATABASE OF SYSTEMATIC REVIEWS*. <https://doi.org/10.1002/14651858.CD011145>
- Eliasberg, W. G. (1953). To examine testamentary and testimonial capacity. *Journal of Criminal Law, Criminology & Police Science*, 44(3), 320–329.
- Faustman, W. O., Moses, J. A., Jr., & Csernansky, J. G. (1990). Limitations of the Mini-mental state examination in predicting neuropsychological functioning in a psychiatric sample. *Acta Psychiatrica Scandinavica*, 81(2), 126–131.
- Fontana, P., Dagnino, F., Cocito, L., et al. (2008). Handwriting as a gauge of cognitive status: A novel forensic tool for posthumous evaluation of testamentary capacity. *Neurological Sciences*, 29(4), 257–261.
- Gerstenecker, A., Martin, R. C., Hebert, K., et al. (2022). Cognitive correlates of impaired testamentary capacity in Alzheimer's dementia. *Archives of Clinical Neuropsychology*, 37(6), 1148–1157.
- Heinik, J., Werner, P., & Lin, R. (1999). How do cognitively impaired elderly patients define "testament": Reliability and validity of the testament definition scale. *The Israel Journal of Psychiatry and Related Sciences*, 36(1), 23–28.
- Kennedy, K. M. (2012). Testamentary capacity: A practical guide to assessment of ability to make a valid will. *Journal of Forensic and Legal Medicine*, 19(4), 191–195.
- Larner, A. J. (2012). Screening utility of the montreal cognitive assessment (MoCA): in place of—or as well as—the MMSE? *International Psychogeriatrics*, 24(3), 391–396.
- Lichtenberg, P. A., Gross, E., & Campbell, R. (2020). A short form of the Lichtenberg financial decision rating scale. *Clinical Gerontologist*, 43(3), 256–265.
- Marson, D. C., Huthwaite, J. S., & Herbert, K. (2004). Testamentary capacity and undue influence in the elderly: A jurispudent therapy perspective. *Law & Psychol. Rev.*, 28, 71.
- Martin, R. C., Gerstenecker, A., Hebert, K., et al. (2022). Assessment of testamentary capacity in older adults: Description and initial validation of a standardized interview instrument. *Archives of Clinical Neuropsychology*, 37(6), 1133–1147.
- Millan, M. J., Agid, Y., Brüne, M., et al. (2012). Cognitive dysfunction in psychiatric disorders: Characteristics, causes and the quest for improved therapy. *Nature Reviews. Drug Discovery*, 11(2), 141–168.
- Molloy, D. W., & Standish, T. I. (1997). A guide to the standardized Mini-mental state examination. *International Psychogeriatrics*, 9(Suppl. 1), 87–94 (discussion 143–150).
- Papageorgiou, S. G., Voskou, P., Economou, A., et al. (2018). Testamentary capacity assessment tool (TCAT): A brief instrument for patients with dementia. *Journal of Alzheimer's Disease*, 61(3), 985–994.
- Peisah, C., Finkel, S., Shulman, K., et al. (2009). The wills of older people: Risk factors for undue influence. *International Psychogeriatrics*, 21(1), 7–15.
- Purser, K., & Lonie, J. (2019). Mapping dementia and cognitive decline in testamentary capacity. *International Journal of Law and Psychiatry*, 66, Article 101450.
- Purser, K., Magner, E. S., & Madison, J. (2015). A therapeutic approach to assessing legal capacity in Australia. *International Journal of Law and Psychiatry*, 38, 18–28.
- Roked, F., & Patel, A. (2008). Which aspects of cognitive function are best associated with testamentary capacity in patients with Alzheimer's disease? *International Journal of Geriatric Psychiatry*, 23(5), 552–553.
- Shulman, K. I., Cohen, C. A., & Hull, I. (2005). Psychiatric issues in retrospective challenges of testamentary capacity. *International Journal of Geriatric Psychiatry*, 20(1), 63–69.
- Shulman, K. I., Himel, S., Hull, I. M., et al. (2017). *Banks v GOODFELLOW (1870): Time to update the test for testamentary capacity*.
- Shulman, K. I., Peisah, C., Jacoby, R., et al. (2009). Contemporaneous assessment of testamentary capacity. *International Psychogeriatrics*, 21(3), 433–439.
- Shulman, K. I., Carole, A., Felice, C., Kirsh, L. B., et al. (2007). Assessment of testamentary capacity and vulnerability to undue influence. *American Journal of Psychiatry*, 164(5), 722–727.
- Sousa, L. B., Simões, M. R., Firmino, H., et al. (2014). Financial and testamentary capacity evaluations: Procedures and assessment instruments underneath a functional approach. *International Psychogeriatrics*, 26(2), 217–228.
- Spiers v Diane Hewston. (2023). EWHC 1145. [https://www.mentalhealthlaw.co.uk/Baker_v_Hewston_\(2023\)_EWHC_1145_\(Ch\)](https://www.mentalhealthlaw.co.uk/Baker_v_Hewston_(2023)_EWHC_1145_(Ch)).

Stankowski, J., & Noffsinger, S. G. (2005). Undue influence. *Journal of the American Academy of Psychiatry & the Law*, 33(2), 268–270.

Sturman, E. D. (2005). The capacity to consent to treatment and research: A review of standardized assessment tools. *Clinical Psychology Review*, 25(7), 954–974.

Wood, S., & Lichtenberg, P. A. (2017). Financial capacity and financial exploitation of older adults: Research findings, policy recommendations and clinical implications. *Clinical Gerontologist*, 40(1), 3–13.