

Research and Evaluation

Avoiding Theoretical Stagnation: A Systematic Review and Framework for Measuring Public Value

Nicholas Faulkner 
Monash University

Stefan Kaufman
Environment Protection Authority Victoria

Public value theory has become a hot topic in public administration research, but its proponents have long recognised difficulties in empirically testing the theory's central propositions. There has been a lack of clarity about how to measure the extent to which organisations are generating public value, which has rendered researchers unable to quantitatively study the causes, consequences and correlates of public value. The current study systematically reviews the growing literature on public value measurement to identify, evaluate, and synthesise available measures. Through a qualitative synthesis of the themes present in published measures, we identify four key components for measuring public value that appear to be important across a range of policy and national contexts. Our review identifies a promising framework that could be used to structure a comprehensive measure of public value and, in doing so, provides a means to progress theoretical development and testing of the public value approach.

Key words: *public value, measurement, systematic review, performance measurement*

Mark Moore's (1995) public value approach to public management, which posits that the task of public sector managers is to create public value, has become a hot topic among public administration academics and practitioners (Bracci et al. 2014; O'Flynn 2007; Rutgers 2015; Stoker 2006). Public value has been described as an idea that has 'made it' in public administration practice and research (Meynhardt and Bartholomes 2011; Talbot 2009). However, despite the substantial attention given to public value by academics and practitioners, there remains a lack of clarity about how to measure the extent to which an organisation has created public value (Marcon 2014; Mendel and Brudney 2014; Meynhardt and Bartholomes 2011). According to Mendel and Brudney (2014:33),

measurement of public value 'remains elusive, with little attention and some speculation'. This lack of clear measurement options has persisted despite researchers' repeated recognition of the need for such measures (Horner and Hutton 2011; Mendel and Brudney 2014; Meynhardt and Bartholomes 2011; Moore 1995; Talbot and Wiggan 2010), and despite calls for more empirical studies on public value (Alford and Hughes 2008; Williams and Shearer 2011). The lack of valid and reliable measures makes it impossible for researchers to quantitatively test hypotheses about the causes and consequences of public value, which puts public value research at high risk of theoretical stagnation (Williams and Shearer 2011), and makes it impossible for practitioners to measure the extent to which their organisations are creating public value.

Accepted for publication: March 20, 2017

Although no single, commonly used measure of public value exists, an increasing number of studies have attempted to develop public value measurement techniques. Here we systematically review published, peer-reviewed research on how to measure the extent to which an organisation is creating public value. Our review makes several empirical, theoretical, and practical contributions to the study of public value. At an empirical level, our systematic review identifies the broad array of constructs that existing measures have included as indicators of public value, and uncovers key gaps in existing research by showing how often such research has used particular research methods and investigated particular policy areas and national contexts. Our review makes a substantial contribution to the theory of public value by identifying four key dimensions that constitute public value across a broad array of national and policy constructs. In contrast to common claims that what constitutes public value differs across contexts (e.g. Benington 2011; Moore 1995; Spano 2014:367), these constructs represent a clear operational meaning of public value that applies across a broad array of national and policy contexts. Finally, at a practical level, our review provides academics and practitioners with a short and accessible overview of existing public value measures. Accordingly, this review provides a ‘state of knowledge’ summary for practitioners who want to measure the extent to which their organisations are generating public value. For researchers and theoreticians, it provides a synthesis of current operationalisation of the theory, which can be evaluated for its consistency and completeness with the theoretical narrative on public value, and serve as a basis to further develop empirical measures and research techniques.

Defining Public Value

Public value is a multifaceted concept that has been approached in several ways (Alford and O’Flynn 2009; Bozeman 2009; Horner et al. 2006; Rutgers 2015; Williams and Shearer 2011) and has attracted criticism for being poorly defined (Prebble 2012; Rhodes and Wanna 2007). At a broad level, public value

has been described as ‘... a comprehensive approach to thinking about public management and about continuous improvement in public services’ (Constable et al. 2008; Moore 1995). Public value is related to, but distinct from, research on public values (Nabatchi 2011; Van der Wal et al. 2015). Public value refers to ‘the value created by government through services, laws regulation and other actions’ (Kelly et al. 2002).¹ It is produced by public managers successfully navigating a strategic triangle (Moore 1995) encompassing (1) producing valued outcomes, and doing so within the constraints of (2) available resources and capability, and (3) the authorising environment of formal and informal jurisdiction, legal frameworks, and mandate. According to Moore (1995:28), ‘the aim of managerial work in the public sector is to create *public* value just as the aim of managerial work in the private sector is to create *private* value’. In contrast, public values refer to normative personal judgements about the ‘social standards, principles, and ideals to be pursued and upheld by government agents and officials’ (Bozeman 2007; Nabatchi 2011). The current article focusses on Moore’s conception of public value (Moore 1995).

The Importance of Measuring Public Value

Measuring the extent to which government bodies are creating public value is important for both practical and scholarly reasons. For practitioners, measuring public value is important for at least three reasons: ‘(1) to meet demands for external accountability; (2) to establish a clear, significant mission and goal for the organisation and (3) to foster a strong sense of internal accountability’ (Moore 2007:97; see also Spano 2014). Measuring public value forces public administrators to be explicit about the types of public value they are seeking to create, which can lead to increased performance (Moore 2007). For scholars, measuring public value is essential for testing hypotheses about the possible causes and consequences of public value. Without an ability to reliably and validly measure an organisation’s public value, it is impossible to quantitatively test hypotheses about how to maximise public value, or

the impact public value has on citizens' lives. Furthermore, without an ability to test hypotheses about public value, theoretical development will remain at risk of stagnation (Williams and Shearer 2011) because researchers will be unable to identify the causes, correlates, and consequences of public value. Williams and Shearer's (2011) systematic review of research on public value highlighted the need for empirical research to evaluate the claims made by both proponents and critics of public value. Without improving the empirical foundation of public value research, 'public value is likely to . . . fall short of offering a broader theory of public enterprise and organization' (Williams and Shearer 2011:1381).

Current Research

Although there have been reviews of public value literature generally (Williams and Shearer 2011), no review has focussed specifically on the measurement of public value, despite its importance, and despite the growing amount of research on this topic. Accordingly, we focus specifically on the measurement of public value. We review how existing research has recommended measuring public value and, through a qualitative synthesis of themes in these measures, we identify the current state of knowledge regarding key components considered necessary for measuring the extent to which an organisation has created public value.

Method

We used the systematic review method to identify and summarise research on public value measurement (Cooper 2016; Petticrew and Roberts 2008). The systematic review method allows large bodies of literature to be identified and synthesised in a transparent and reproducible manner. Systematic reviews differ from traditional, narrative literature reviews in several ways. Narrative reviews typically 'do not involve a systematic search of the literature, . . . often focus on a subset of studies in an area chosen based on availability or author selection' (Uman 2011:57), and typically do not

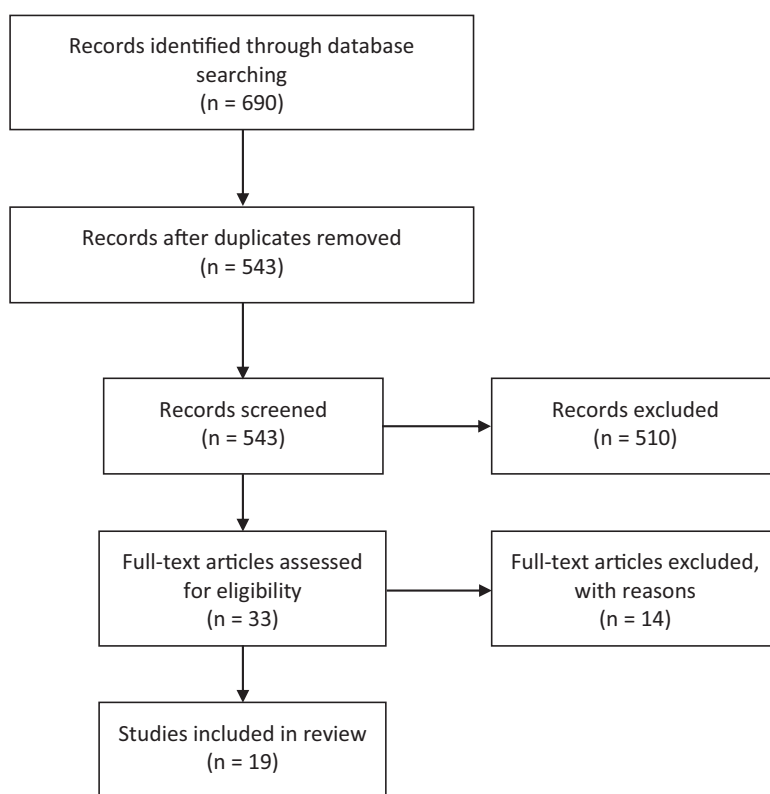
include quality assessments of included studies (Grant and Booth 2009). Although narrative reviews are often useful, they may be based on a biased selection of studies, and they may struggle to reconcile findings from studies with conflicting results (Uman 2011). In contrast to narrative literature reviews, systematic reviews have clear and explicit search strategies, inclusion criteria, synthesis techniques, and often include quality assessments of included studies. Systematic reviews have the benefit of being transparent about inclusion and exclusion decisions, and generally more comprehensive than non-systematic reviews. A disadvantage, however, is that systematic reviews take longer to complete than narrative reviews (Grant and Booth 2009). Additionally, like most narrative reviews, systematic reviews also often rely on electronic database keyword indexing. In fields where their use is common, systematic reviews are treated as a 'fundamental scientific activity' that is distinct from and vastly superior to narrative reviews (Mulrow 1994:597).

Search Strategy

To identify published and peer-reviewed research on the measurement of public value, we searched the following databases: Web of Science, Proquest, Business Source Complete, Emerald, PAIS International, Worldwide Political Science Abstracts, and Econlit. The databases were searched using the term 'public value*', combined with the terms 'measur*', 'scale*', 'metric*', or 'checklist'. Search terms and database choices were developed in consultation with a university librarian before the final search process commenced. All searches were restricted to include only studies published during or after 1995, which is the year that Moore's (1995) first book on public value was published. The last search was conducted on 27 January 2016.

Inclusion Criteria

Studies were included in this review if they met all of the following criteria.

Figure 1. Flow Diagram Showing the Search and Screening Process

- Topic: Studies must propose a method by which government bodies can measure the extent to which they are generating ‘public value’, as conceptualised by Moore (1995). Studies which solely used contingent valuation or willingness to pay methods without considering Moore’s broader conception of public value (e.g. Kwak and Yoo 2012) were thus excluded, as were studies which sought to identify what citizens value, rather than measure the extent to which an organisation or policy has created public value (e.g. Jørgensen and Bozeman 2007; Karunasena and Deng 2011).
- Study design: All quantitative, qualitative, and conceptual studies were included.
- Language: Only English language studies were included.
- Year of publication: January 1995 to January 2016.

- Publication status: Published in peer-reviewed journal articles or books. We used this criterion to ensure that all included articles had passed the minimum peer-review quality hurdle.

Study Selection and Data Extraction

All studies identified in the database searches (n = 543) were first subjected to title and abstract screening. At this stage, we examined titles and abstracts and retained only articles that appeared to meet all inclusion criteria (n = 33). At the next stage, we examined full text articles, and included 19 studies that met all inclusion criteria. The search process is reported in Figure 1. For each included study, we extracted information on study characteristics, methodology, context, populations, and dimensions of public value identified as important for measurement.

Table 1. Quality assessment of quantitative measurement studies

Criteria	Karkin and Janssen (2014)	Meynhardt and Bartholomes (2011)	Al-Hujran et al. (2015)
Internal consistency: Was internal consistency assessed?	No	Yes	Yes
Reliability: Was test–retest or inter-rater reliability assessed?	No	No	No
Content validity: Was an assessment of whether all items refer to relevant aspects of the construct to be measured performed?	Yes	Yes	Unsure
Structural validity: Was the factor structure or dimensionality of the items assessed?	No	Yes	Yes
Construct validity: Was the measure shown to correlate with theoretically related measures?	No	No	Yes
Responsiveness: Was the measure shown to change over-time in response to changes in other variables?	No	No	No
Total score (out of 6)	1	3	3
Overall quality rating	Low	Moderate	Moderate

Quality Appraisal

Research on systematic review methodology generally recommends that the quality of included studies is assessed and used to interpret results (Shea et al. 2007). Quality assessments are usually conducted using standardised checklists (Shea et al. 2007; Verhagen et al. 1998; Whiting et al. 2003). Although quality assessment criteria exist that are appropriate for quantitative measurement articles (i.e. studies that develop and validate quantitative measures; e.g. Mokkink et al. 2010), no commonly accepted criteria exist for qualitative or conceptual measurement articles (i.e. studies that provide suggestions for measurement frameworks without developing or quantitatively validating particular measures; see Thomas and Harden 2008). Accordingly, we assessed the quality of quantitative measurement articles using a short checklist that contained items adapted from both measurement quality assessment criteria used in other disciplines (Mokkink et al. 2010), and common recommendations for scale development (Clark and Watson 1995; DeVellis 2012), but did not assess the quality of conceptual or qualitative studies. Of the three quantitative studies included in this review, one was rated as low qual-

ity, and the other two were rated as moderate quality (see Table 1).

Synthesis Approach

Thematic synthesis (Thomas and Harden 2008) was used to identify overarching analytical themes that captured the proposed public value measurement dimensions identified in the reviewed studies. Thematic synthesis is a method to identify qualitative themes across studies included in a systematic review. We employed the approach by, first, extracting the public value dimensions identified in each study. We then identified overarching analytical themes that included or described all or almost all of the public value dimensions extracted from each study.

Results

Publishers and Countries

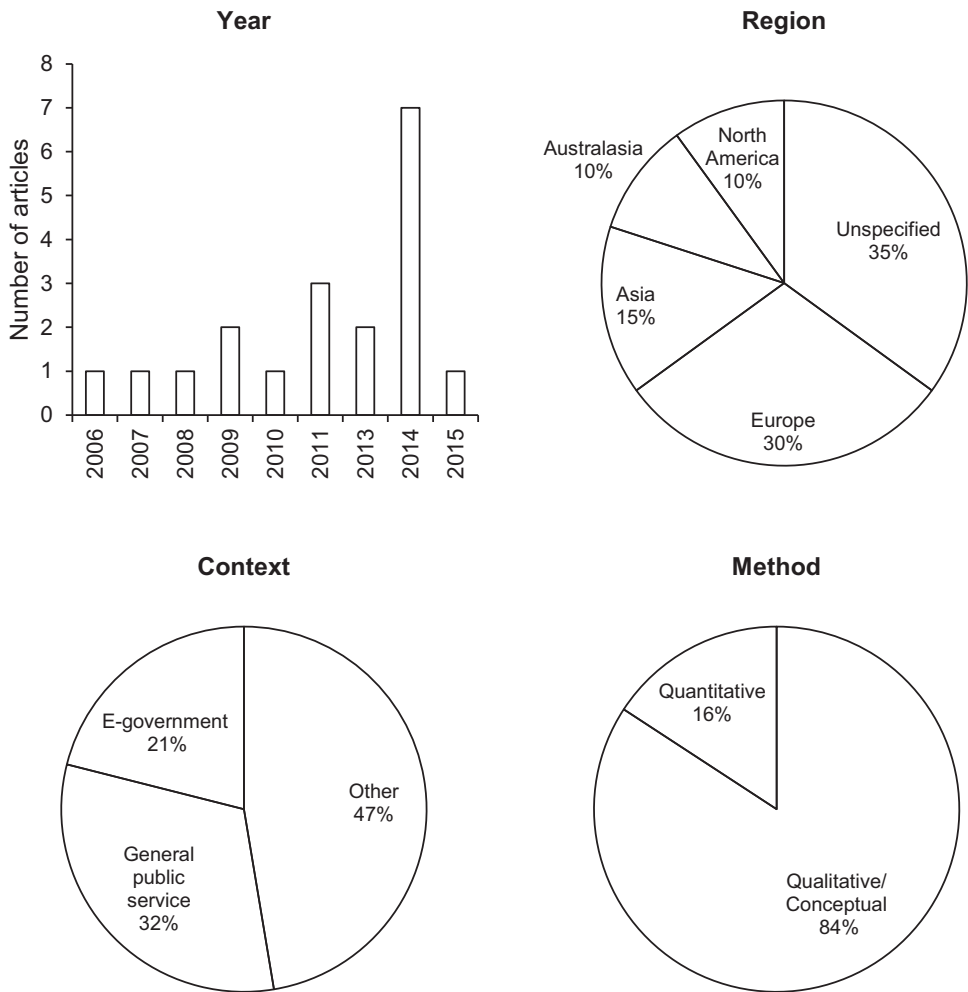
Table 2 shows the background characteristics of each study included in this review. Of these studies, approximately half ($n = 10$, 52.6%) were journal articles, whereas the remaining half were book chapters ($n = 8$, 42.1%) or books ($n = 1$, 0.5%). The journal articles

Table 2. Characteristics of included studies

Reference	Publication type	Context	Country	Methodology
Al-Hujran et al. (2015)	Journal article	e-Government	Jordan	Quantitative
Benington (2009)	Journal article	General public service	Unspecified	Qualitative/conceptual
Benington (2011)	Book chapter	General public service	Unspecified	Qualitative/conceptual
Bozeman et al. (2015)	Journal article	Technology transfer programs and policies	Unspecified	Qualitative/conceptual
Bracci et al. (2014)	Book chapter	Publicly owned theatre	Italy	Qualitative/conceptual
Brookes and Wiggan (2009)	Journal article	Sport services	United Kingdom	Qualitative/conceptual
Christensen et al. (2006)	Journal article	Property subdivision policies	United States	Qualitative/conceptual
Collins (2007)	Journal article	Public service broadcasting	United Kingdom	Qualitative/conceptual
Conolly (2013)	Book chapter	Museums	Australia	Qualitative/conceptual
Heeks (2008)	Book chapter	e-Government	Unspecified	Qualitative/conceptual
Karkin and Janssen (2014)	Journal article	e-Government	Turkey	Quantitative
Karunasena and Deng (2011)	Journal article	e-Government	Sri Lanka	Qualitative/conceptual
Liguori et al. (2014)	Book chapter	General public service	Italy	Qualitative/conceptual
Marcon (2014)	Book chapter	General public service	Unspecified	Qualitative/conceptual
Meynhardt and Bartholomes (2011)	Journal article	Federal Labour Agency	Germany	Quantitative
Modugno et al. (2014)	Book chapter	Universities	Italy	Qualitative/conceptual
Moore (2013)	Book	General public service	Unspecified	Qualitative/conceptual
Spano (2014)	Book chapter	General public service	Unspecified	Qualitative/conceptual
Talbot and Wiggan (2010)	Journal article	Supreme audit institutions	Australia, Canada, United States, and New Zealand	Qualitative/conceptual

(n = 10) were published in a variety of public administration (e.g. *International Journal of Public Administration*), public management (e.g. *Public Management Review*), and other journals (e.g. *International Journal of Information Management*). Books and book chapters (n = 9) were published by Harvard University Press, Emerald, and Palgrave MacMillan.

The studies were conducted in a range of nations across several contexts. Roughly one quarter (n = 5, 26.3%) of the included studies were conducted in Anglosphere nations, and roughly a third (n = 7, 36.8%) did not specify a particular national context. Several studies were conducted outside of Anglosphere nations: three (21.1%) studies were conducted in

Figure 2. Summary of Study Background Characteristics

Italy: one in Germany, one in Jordan, and one in Turkey. Studies were also conducted across a wide variety of public policy contexts. Aside from ‘general public service’ ($n = 6$, 31.6%), the only other context with more than one study was ‘e-government’ ($n = 4$, 21.1%). Study background characteristics are summarised in Figure 2.

Research Methods

The research methods employed by included studies were overwhelmingly qualitative or conceptual ($n = 16$, 84.2%), rather than quantitative ($n = 3$, 15.8%). Qualitative/conceptual

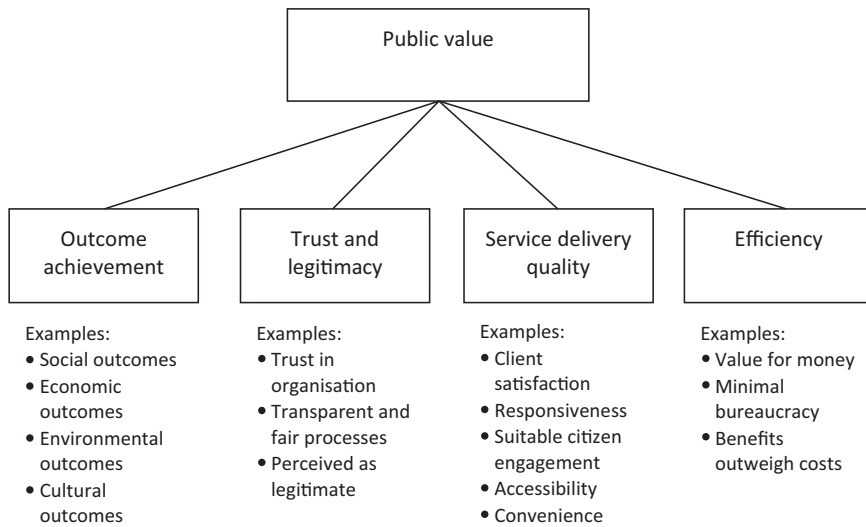
studies primarily aimed to identify measurement frameworks, or illustrate the application of a measurement framework to a particular context, but did not develop or validate quantitative measures. In contrast, quantitative studies attempted to develop and validate quantitative measure of public value, or aspects of public value.

Dimensions of Public Value

The included studies identified a broad array of public value dimensions. The dimensions identified in each study are shown in detail in Table 3. Four analytical themes were identified

Table 3. Public value measurement dimensions identified in each study

Reference	Public value measurement dimensions identified
Al-Hujran et al. (2015)	Unidimensional scale that included items relating to efficiency, valuing the service, transparency, accountability
Benington (2009)	Public satisfaction; improved ecological, political, economic, and social outcomes; improved efficiency and efficacy; co-creation
Benington (2011)	Public satisfaction, economic value (generating economic activity/employment), social and cultural value (social capital/cohesion), political value (democratic dialogue, public participation), ecological value (sustainable development, reducing pollution, waste, global warming)
Bozeman et al. (2015)	Improved publicly valued outcomes in a wide variety of areas (e.g. national security, food safety, human nutrition, human health, environmental/natural resource protection, increased access to knowledge)
Bracci et al. (2014)	Social value (from user perspective), tangible economic value (from administration's perspective), intangible economic value (from administration's perspective)
Brookes and Wiggan (2009)	Process/equity, social outcomes and value for money, resources and efficiency, service quality, trust and legitimacy
Christensen et al. (2006)	Primary dimensions were protecting citizens' rights, improving public health and safety; secondary dimensions were political support, administrative feasibility, efficiency
Collins (2007)	Willingness to pay, reach, impact, quality, value for money
Conolly (2013)	Number of people who use or access service (e.g. stats about number of people visiting museums); meeting citizen expectations about services, cultural and social progress
Heeks (2008)	Service delivery (take-up, satisfaction, information, choice, importance, fairness, cost), outcome achievement, trust in public institutions, efficiency for organisation, efficiency for users, democracy/political values (openness, transparency, participation)
Karkin and Janssen (2014)	Accessibility, citizen engagement, transparency, responsiveness, dialog, balancing of interests
Karunasena and Deng (2011)	Delivery of public services, operating an effective public organisation, achievement of outcomes, development of trust
Liguori et al. (2014)	Financial performance (e.g. revenues, expenditure), non-financial performance (e.g. efficiency, customer satisfaction, service quality and appropriateness)
Marcon (2014)	Service outputs, satisfaction, outcomes, maintaining trust/legitimacy, greater effectiveness orientation, service delivery and system maintenance
Meynhardt and Bartholomes (2011)	Empirical dimensions: institutional performance (service and delivery), moral obligation (improving equal opportunity), political stability (social cohesion, social peace); conceptual dimensions: moral-ethical (treated fairly, equally, and justly), political-social (belongingness, cohesion, solidarity), hedonistic-aesthetical (maximising pleasure, personal safety, and experience of public spaces as cultural achievements), instrumental-utilitarian (functionality, extent to which service effectively solves a problem)
Modugno et al. (2014)	Process, efficiency, output, customer satisfaction, outcome, effectiveness, impact
Moore (2013)	Customer satisfaction (including client and citizen satisfaction), outcomes (including social outcomes), various other outcomes that may be valued by the public in some contexts (e.g. efficiency in use of money, fair use of authority).
Spano (2014)	Extent to which the organisation's outcomes and objectives have been achieved (e.g. outcomes, satisfaction, willingness to pay, citizen participation, trust)
Talbot and Wiggan (2010)	Trust and legitimacy, process and equity, resources and efficiency, social outcomes and value for money, services and users focus

Figure 3. Summary of Proposed Public Value Measurement Dimensions

(see Figure 3): (1) outcome achievement, (2) trust and legitimacy, (3) service delivery quality, and (4) efficiency. These themes represent four key constructs that the reviewed studies suggest are important dimensions of public value measurement. The themes also broadly reflect the most common terms used to describe public value measurement dimensions (as shown in Figure 4).

Outcome Achievement

‘Outcome achievement’ refers to the extent to which a public body is improving publicly valued outcomes across a wide variety of areas. For example, Bozeman et al. (2015) mention several such outcomes, including national security, food safety, human health, environmental protection, and access to knowledge. Benington (2009, 2011) similarly suggests that benefits to economic activity and employment, social capital and cohesion, and environmental outcomes (e.g. reducing pollution, waste, and global warming) could be considered as aspects of public value. Similarly, several authors note that public value can be indicated by improvements in social outcomes (Bracci et al. 2014; Brookes and Wiggan 2009; Conolly 2013; Meynhardt and Bartholomes 2011; Moore 2013; Spano 2014; Talbot and Wiggan 2010).

Studies often imply that the specific types of outcomes achieved by particular agencies will differ. For example, a labour agency may be focussed on reducing unemployment or increasing equal opportunity (e.g. Heeks 2008; Meynhardt and Bartholomes 2011), whereas an environmental regulator may be focussed on reducing pollution. However, to create public value government bodies need to be contributing to improved outcomes in some way (Bozeman et al. 2015; Heeks 2008).

There are only a small number of specific recommendations in the reviewed studies about how to measure outcome achievement. Conolly (2013:119) implies that public museums might measure the extent to which they are assisting in improving individuals’ ‘opportunities to participate in Australia’s cultural life’ by examining museum visitor numbers. In contrast, Meynhardt and Bartholomes (2011:297) use surveys to measure the extent to which a labour agency is perceived to: ‘successfully promote disabled people’s participation in the labor market’, ‘effectively contribute to social cohesion’, and ‘effectively support immigrants’ skills development’. No studies presented a standardised way to measure outcome achievement across different government bodies, nor was there clear guidance about whether outcome achievement should be measured



using objective indicators, such as unemployment or pollution data, or on subjective indicators, such as surveys of citizens' and stakeholders' perceptions.

Trust and Legitimacy

'Trust and legitimacy' refers to the extent to which an organisation and its activities are trusted and perceived to be legitimate by the public and by key stakeholders. This dimension includes the extent to which the public trust the particular institution (e.g. Heeks 2008; Talbot and Wiggan 2010), trust the programs or services delivered by institution (e.g. Heeks 2008; Karunasena and Deng 2011), and perceive an institution to be delivering services transparently and fairly (e.g. Karkin and Janssen 2014; Meynhardt and Bartholomes 2011). Talbot and Wiggan (2010) argue that trust and legitimacy are at the heart of public value creation, largely because organisations that are trusted and perceived to be legitimate are best able to secure the support needed to carry out their objectives.

Several of the reviewed studies suggested that trust and legitimacy may be measured using surveys, but no single, validated measure of trust was presented. Heeks (2008) explicitly states that trust can be measured by user survey, but does not provide any survey measure for this purpose. Meynhardt and Bartholomes (2011) include at least one item in their scale ('My local agency is an institution one can trust') that appears to measure trust, but did not attempt to develop nor validate a measure of trust and legitimacy.

Service Delivery Quality

'Service delivery quality' refers to the extent to which services are experienced as being delivered in high-quality manner that is considerate of users' needs. It is expected to be maximised when individuals who interact with the service are satisfied, and when they perceive the services to be responsive to their needs, accessible, convenient, and incorporate sufficient citizen engagement (see Al-Hujran et al. 2015; Benington 2009, 2011; Brookes and Wiggan 2009; Heeks 2008; Karkin and Janssen 2014; Spano 2014). Although no studies explicitly developed or validated quantitative measures

of service delivery quality, at least two studies provided suggestions that may be useful for developing such a measure. Heeks (2008) suggests that satisfaction, which is expected to be indicative of service delivery quality, could be measured using surveys. Similarly, Meynhardt and Bartholomes' (2011) measure of public value included items related to perceived service quality, customer satisfaction, and responsiveness to feedback. Finally, Al-Hujran et al.'s (2015) five-item public value scale includes an item about the perceived convenience of the service.

Efficiency

'Efficiency' refers to the extent to which an organisation is achieving maximal benefits with minimal resources. It is expected to be high when the benefits provided by an organisation are perceived to outweigh the costs of that organisation (Talbot and Wiggan 2010), when "unnecessary" bureaucracy is avoided (see Meynhardt and Bartholomes 2011), and when an organisation is perceived to offer 'value for money' (Brookes and Wiggan 2009; Collins 2007; Talbot and Wiggan 2010).

How to measure efficiency is not well explained in existing research, despite 'efficiency' or 'value for money' being noted in 9 of the 19 studies as a relevant measurement dimension. Meynhardt and Bartholomes' (2011:297) scale includes an item assessing agreement with the statement 'My local agency acts flexibly and avoids unnecessary bureaucracy', but does not attempt to measure perceived efficiency comprehensively. Moore (2013) presents an accounting approach akin to an 'income statement' whereby costs (including financial and social costs) are presented alongside, and compared with, measures of generated value (including 'mission achievement' and 'client satisfaction'). However, Moore (2013) does not provide a technique to measure how much value has been generated, thereby precluding any simple estimates of value for money.

Quantitative Measures of Public Value

Although all reviewed studies identified dimensions of public value, very few made clear

Table 4. Example items from quantitative measures of public value

Reference	Description of measure	Example item(s)
Al-Hujran et al. (2015)	Survey measure containing four items which loaded on a single dimension ('public value')	'I value the convenience of using e-government portal and/or Ministry's website(s) to access government services'
Meynhardt and Bartholomes (2011)	Survey measure containing 16 items. Items loaded onto three first-order factors ('institutional performance', 'moral obligation', 'political stability'), which in-turn loaded onto a single second-order factor ('public value')	<i>Institutional performance</i> : 'My local agency delivers high-quality service' <i>Moral obligation</i> : 'My local agency provides special support for handicapped people in the labor market' <i>Political stability</i> : 'My local agency delivers an important contribution so that nobody "falls through the cracks"'
Karkin and Janssen (2014)	Detailed checklist containing over 20 items ^a purported to measure six dimensions ('accessibility', 'citizen engagement', 'transparency', 'responsiveness', 'dialog', 'balancing of interests') of public value.	<i>Accessibility</i> : no single items reported <i>Citizen engagement</i> : 'Questionnaire' available? (Yes/no) <i>Transparency</i> : 'Publications of council reports and performance' available? (Yes/no) <i>Responsiveness</i> : 'Time taken to receive a response' <i>Dialog</i> : 'Direct line for phone' (Yes/no) <i>Balancing of interests</i> : 'Activities' available? (Yes/no)

Note: ^aKarkin and Janssen (2014) did not report exactly how many items were included in their complete checklist.

and direct recommendations about how to turn these qualitative dimensions into quantitative scores. Several studies noted difficulties associated with measurement. For example, Talbot and Wiggan (2010:64) noted that some dimensions of public value 'are difficult to quantify but clearly of immense importance'. Despite these difficulties, three studies (Al-Hujran et al. 2015; Karkin and Janssen 2014; Meynhardt and Bartholomes 2011) presented quantitative measures of public value. Two of these measures (Al-Hujran et al. 2015; Meynhardt and Bartholomes 2011) were survey measures, and one (Karkin and Janssen 2014) was a detailed checklist. Of the two survey measures, one (Al-Hujran et al. 2015) measured the perceived public value of e-government services, whereas the other (Meynhardt and Bartholomes 2011) measured the perceived public value of a federal labour agency. Al-Hujran et al.'s (2015) scale included five items which loaded onto a single factor labelled 'perceived public value'. Meynhardt and Bartholomes' (2011) scale included 16 items which loaded onto three first-order factors (labelled 'institutional per-

formance', 'moral obligation', and 'political stability') which, in turn, loaded on a second-order factor labelled 'public value'. Karkin and Janssen's (2014) detailed checklist was designed to measure government websites' public value. The checklist contains over 20 items that assess specific website design features related to public value. Example items from each measure, including subscales where available, are shown in Table 4. None of the quantitative measures were developed in a manner that could be consistently applied across all government contexts. Instead, all measures were specific to the contexts for which they were developed and covered only some dimensions of public value. Additionally, none of the measures were validated to a high standard according to the quality assessment criteria shown in Table 1.

Discussion

This is the first known systematic review of research on the measurement of public value.

Despite the importance of measurement for testing public value theory's propositions, and despite an increasing number of studies being published on this topic, no systematic reviews of this research existed. Our aims in reviewing this research were to (a) summarise how existing research suggests organisations can measure the extent to which they are creating public value, (b) identify dimensions of public value that are consistent across these suggested measurement approaches, and (c) identify areas for future research.

A key finding of our review is that four themes captured almost all the public value dimensions outlined by included studies. The four themes identified were outcome achievement, trust and legitimacy, service delivery quality, and efficiency. These four themes appear to represent key dimensions essential for measuring public value.² Accordingly, we argue that these constructs should be included in future research on the measurement of public value, and considered by practitioners for inclusion in public value measures.

The finding that these four themes encompassed almost all the public value dimensions deemed important across a range of policy and national contexts has substantial implications for future research on public value measurement. Most existing research has assumed that public value measures will need to differ across organisations (e.g. Hills and Sullivan 2006; Spano 2014). This assumption appears to be based on the belief that different agencies will create different types of public value that need to be measured differently. A limitation of measuring public value in this way, however, is that it precludes comparisons of the public value created by different organisations. Additionally, it ignores the possibility that public value may be capable of being measured in a largely standardised manner across organisations. This might be particularly valuable when assessing the public value contributed by different agencies to whole of government policy priorities, and collaborative initiatives involving multiple government, civil, and private actors. To extend Moore's (1995) analogy, if private shareholders are best served by ensuring that organisations achieve above-average performance,

established through performance benchmarking against 'like' organisations, public managers could also benefit from the ability to benchmark their performance in terms of public value.

The four dimensions of public value identified in the current review seem to be applicable across most or all types of public organisations. Similarly, although much research notes that organisations need to 'focus on what the public values' (Benington 2011; Spano 2014:367), the themes identified in this review suggest a set of general dimensions that may be valued by most societies in most policy contexts. Accordingly, the findings of this review suggesting that it may be possible to develop a universal measure of public value that could be used across a wide variety of policy and national contexts.

Although the development of such a measure needs to be the subject of future research, several considerations should inform this scale development. A public value measure would be more valuable if it allowed comparisons and benchmarking across organisations, meaning that the measure would need to be sufficiently general to be used consistently across contexts. However, a disadvantage of using a consistent measure may be that such a measure could not capture all the subtle nuances of value creation that differ across contexts. Additionally, a consistent measure may be incapable of capturing changes in citizens' perceptions about what matters at a given time and place. Accordingly, researchers developing measures of public value should aim to use techniques to maximise the generalisability of the measure while balancing needs to measure aspects of value that are context-specific. One approach might be to use a modular approach in which both core generic measures (which, as our review suggests, are likely to be largely invariant), plus specific measures that change from context to context are included. The generic measures could be used for cross-organisational comparisons, while the specific measures could address the nuances of the specific context.

The included studies made recommendations about what constructs should be included in

measures of public value, but typically did not specify how to measure those constructs. All but 3 of the 19 studies used qualitative or conceptual approaches that listed indicators purported to be important for measuring public value, but did not develop or validate ways to quantitatively measure performance on those indicators. Of the three studies that reported a quantitative measure of public value, none provided a measure of public value that either: captured all four of the themes identified in the current review, or was applicable across more than one policy context. However, some of the items included in these quantitative studies could be built upon to develop a more comprehensive, widely applicable, and better validated measure of public value. Another limitation of the three quantitative studies was that the quality of the validation of these measures was rated as moderate at best. None of the studies reported measures of test-retest or inter-rater reliability (which examine the extent to which the measure produces consistent scores across time or raters), and tests for construct validity (which examine the extent to which a measure seems to measure the construct it is purported to measure) were not conducted in two of the three quantitative studies.

The findings of the review indicate several areas where further research is needed. First, our findings highlight the need for more quantitative research to develop and validate measures of public value. Additionally, future quantitative research needs a stronger focus on assessing construct validity and test-retest reliability of the measures under investigation. Second, building on this point, our findings indicate that future research on the measurement of public value should include items that capture the four key dimensions of public value identified in this review. Although the four dimensions identified here need to be subjected to empirical testing to examine the extent to which they are quantitatively distinct, they provide a framework that can be used to guide the development and validation of public value measures. Given that the subject of some of the four dimensions are themselves the subject of focussed research – for example, there is al-

ready research on trust in government (Döring 1992; Faulkner et al. 2015; Feldman 1983) and service quality (Parasuraman et al. 1988, 1991) – it is somewhat surprising that public value measurement efforts to date have not transparently attempted to transfer validated measures from these fields. Additionally, the four dimensions are also core topics in government performance reporting and performance measurement ‘grey literature’ (Productivity Commission Australia 2016). Testing the relevance, and inter-operability, of measures of the individual constructs from other fields of research and existing practice is a promising area of inquiry.

Although our review provides the most comprehensive overview of published approaches to public value measurement, it is not without limitations. First, it relies on the completeness of the databases searched. We searched a broad variety of databases using search terms developed in conjunction with a specialist librarian, but it is possible that there may have been relevant articles that were not indexed on any of the databases searched. However, given we searched both major (e.g. Web of Science, Proquest) and specialised databases (e.g. PAIS International, Worldwide Political Science Abstracts), we assessed this as low risk. Second, because our measurement framework is based on a synthesis of themes in existing literature, it is only as complete as the literature as a whole. If all the included studies failed to include an aspect of public value, that aspect will also be missing from the synthesised measurement framework presented here. Our framework does not depend on any one piece of literature being based on a strong understanding of PV concepts, but does assume that (a) each piece of literature has at least a partial understanding of PV concepts, and (b) as a whole, the literature reviewed includes all relevant public value concepts. Given our review includes a broad range of public value measurement literature – including a measurement approach developed by the original public value theoretician (Moore 2013) – it is unlikely that our framework is missing important public value dimensions, but this remains a possibility.

Conclusion

This systematic review provides academics and practitioners with an accessible summary of the growing peer-reviewed literature on public value measurement. The review contributes to public value research by identifying and defining four central constructs important for measuring public value. This is valuable for practitioners who wish to evaluate their performance via a public value framework in transparent, comprehensive and comparable way, and valuable for researchers interested in improving the empirical basis of public value research (Williams and Shearer 2011). It is also critical for the development of public value theory. Our synthesis of the current state of empirical operationalisation of public value provides theoreticians with an insight into how well and completely the theory is represented in current research. Continued dialogue between theory, empirical research, and practice is critical because without valid and reliable measures of public value theoretical development is likely to stagnate.

Our review shows that there are four widely applicable, but still coherent, dimensions public value concept that appear to resonate worldwide and across diverse contexts. These four dimensions provide a framework that can be used to develop a new, widely applicable measure of public value. In this way, our review provides a promising avenue through which theoretical development can proceed on the basis of quantitative evidence.

An additional contribution of the current review is to identify gaps in existing literature on public value measurement, and highlight key future research needs. Despite the growing number of studies in this area, very few studies have attempted to develop and validate quantitative measures of public value. Without valid, reliable, and widely applicable measures, it will remain impossible to test hypotheses about the causes and consequences of public value, and impossible for organisations to reliably measure the extent to which they are generating public value. As such, progressing to empirical testing of the four dimensions identified in this review is an important next step for practition-

ers, researchers, and theoreticians who wish to see public value research mature.

Endnotes

1. Several scholars have argued that public value can be created not only by government bodies, but also by private organisations (e.g. Benington 2011). We take no position on this debate, but note that all articles meeting the current review's inclusion criteria focused on government or primarily government-run (e.g. state universities) organisations. Accordingly, more research is needed to assess the extent to which our proposed framework is relevant for measuring public value created by private organisations.

2. These dimensions are also consistent with those identified by Moore (2013) on his 'public value scorecard', which is perhaps unsurprising given that Moore (2013) was one of the records included in this review. The dimensions we outline capture all of those stated in the 'public value account' section of Moore's (2013) scorecard, albeit with some differences in labelling, groupings, and emphasis. For example, Moore (2013) includes 'social outcomes' and 'mission achievement' (which is similar to 'outcome achievement' in our model), 'client satisfaction' (which is an element of 'service delivery quality' in our model), and 'justice and fairness' (which is part of 'trust and legitimacy' in our model). Our framework has the benefit of being informed by a larger number of studies and contexts than those included by Moore (2013).

References

- Alford, J. and O. Hughes. 2008. 'Public Value Pragmatism as the Next Phase of Public Management'. *The American Review of Public Administration* 38(2):130–148.
- Alford, J. and J. O'Flynn. 2009. 'Making Sense of Public Value: Concepts, Critiques and Emergent Meanings'. *International Journal of Public Administration* 32(3–4):171–191.
- Al-Hujran, O., M. M. Al-Debei, A. Chatfield and M. Migdadi. 2015. 'The Imperative of

- Influencing Citizen Attitude Toward e-Government Adoption and Use'. *Computers in Human Behavior* 53(1):189–203.
- Benington, J. 2009. 'Creating the Public in Order to Create Public Value?' *International Journal of Public Administration* 32(3):232–249.
- Benington, J. 2011. 'From Private Choice to Public Value?' In J. Benington and M. H. Moore (eds.), *Public Value: Theory and Practice* (pp. 31–51). Basingstoke: Palgrave Macmillan.
- Bozeman, B. 2007. *Public Values and Public Interest: Counterbalancing Economic Individualism*. Washington, DC: Georgetown University Press.
- Bozeman, B. 2009. 'Public Values Theory: Three Big Questions'. *International Journal of Public Policy* 4(5):369–375.
- Bozeman, B., H. Rimes and J. Youtie. 2015. 'The Evolving State-of-the-Art in Technology Transfer Research: Revisiting the Contingent Effectiveness Model'. *Research Policy* 44(1):34–49.
- Bracci, E., E. D. Gagliardo and M. Bigoni. 2014. 'Performance Management Systems and Public Value: A Case Study'. In J. Guthrie, G. Marcon, S. Russo and F. Farneti (eds.), *Public Value Management, Measurement and Reporting* (pp. 129–157). Bingley: Emerald.
- Brookes, S. and J. Wiggan. 2009. 'Reflecting the Public Value of Sport'. *Public Management Review* 11(4):401–420.
- Christensen, P., S. Haskins, J. Hogue and J. Koldoff. 2006. 'Lot Splitting and Development Regulation: The Information Asymmetries and Free Rider Issues Associated with Arizona's Wildcat Development'. *Perspectives in Public Affairs* 3(1):37–59.
- Clark, L. A. and D. Watson. 1995. 'Constructing Validity: Basic Issues in Objective Scale Development'. *Psychological Assessment* 7(3):309–319.
- Collins, R. 2007. 'The BBC and "Public Value"'. *Medien & Kommunikationswissenschaft* 55(2):164–184.
- Conolly, L. 2013. 'Measuring Public Value: Reflections from Australia'. In C. A. Scott (ed.), *Museums and Public Value: Creating Sustainable Futures* (pp. 113–127). London: Routledge.
- Constable, S., E. Passmore and D. Coats. 2008. *Public Value and Local Accountability in the NHS*. London: Work Foundation.
- Cooper, H. 2016. *Research Synthesis and Meta-Analysis: A Step-by-Step Approach*. Thousand Oaks, CA: SAGE Publications.
- DeVellis, R. F. 2012. *Scale Development: Theory and Applications*. Thousand Oaks, CA: SAGE Publications.
- Döring, H. 1992. 'Higher Education and Confidence in Institutions: A Secondary Analysis of the "European Values Survey," 1981–83'. *West European Politics* 15(2):126–146.
- Faulkner, N., A. Martin and K. Peyton. 2015. 'Priming Political Trust: Evidence from an Experiment'. *Australian Journal of Political Science* 50(1):164–173.
- Feldman, S. 1983. 'The Measurement and Meaning of Trust in Government'. *Political Methodology* 9(3):341–354.
- Grant, M. J. and A. Booth. 2009. 'A Typology of Reviews: An Analysis of 14 Review Types and Associated Methodologies'. *Health Information & Libraries Journal* 26(2):91–108.
- Heeks, R. 2008. 'Benchmarking e-Government: Improving the National and International Measurement, Evaluation and Comparison of e-Government'. In Z. Irani and P. Love (eds.), *Evaluating Information Systems: Public and Private Sector* (pp. 257–301). Oxford: Elsevier.
- Hills, D. and F. Sullivan. 2006. *Measuring Public Value 2: Practical Approaches*. London: Work Foundation.
- Horner, L. and W. Hutton. 2011. 'Public Value, Deliberative Democracy and the Role of Public Managers'. In J. Benington and M. H. Moore (eds.), *Public Value: Theory and Practice* (pp. 112–127). New York: Palgrave Macmillan.
- Horner, L., R. Lekhi and R. Blaug. 2006. *Deliberative Democracy and the Role of Public Managers*. London: The Work Foundation.
- Jørgensen, T. B. and B. Bozeman. 2007. 'Public Values an Inventory'. *Administration & Society* 39(3):354–381.
- Karkin, N. and M. Janssen. 2014. 'Evaluating Websites from a Public Value Perspective: A Review of Turkish Local Government Websites'. *International Journal of Information Management* 34(3):351–363.
- Karunasena, K. and H. Deng. 2011. 'Critical Factors for Evaluating the Public Value of e-Government in Sri Lanka'. *Government Information Quarterly* 29(1):76–84.
- Kelly, G., G. Mulgan and S. Muers. 2002. *Creating Public Value: An Analytical Framework for Public Service Reform*. London: UK Cabinet Office.
- Kwak, S.-Y. and S.-H. Yoo. 2012. 'The Public Value of a National Library: Results of a Contingent Valuation Survey'. *Journal of Librarianship and Information Science* 44(4):263–271.
- Liguori, M., M. Sicilia and I. Steccolini. 2014. 'Public Value as Performance: Politicians' and Managers' Perspectives on the Importance Of

- Budgetary, Accruals and Non-Financial Information'. In J. Guthrie, G. Marcon, S. Russo and F. Farneti (eds.), *Public Value Management, Measurement and Reporting* (pp. 85–104). Bingley: Emerald.
- Marcon, G. 2014. 'Public Value Theory in the Context Of Public Sector Modernization'. In J. Guthrie, G. Marcon, S. Russo and F. Farneti (eds.), *Public Value Management, Measurement and Reporting* (pp. 323–351). Bingley: Emerald.
- Mendel, S. C. and J. L. Brudney. 2014. 'Doing Good, Public Good, and Public Value: Why the Differences Matter'. *Nonprofit Management & Leadership* 25(1):23–40.
- Meynhardt, T. and S. Bartholomes. 2011. '(De)Composing Public Value: In Search of Basic Dimensions and Common Ground'. *International Public Management Journal* 14(3):284–308.
- Modugno, G., G. Curiel and G. Ventin. 2014. 'Conceptual Framework and Empirical Evidence of Public Value: The Case of the Italian Higher Education Sector'. In J. Guthrie, G. Marcon, S. Russo and F. Farneti (eds.), *Public Value Management, Measurement and Reporting* (pp. 189–224). Bingley: Emerald.
- Mokkink, L. B., C. B. Terwee, D. L. Patrick, J. Alonso, P. W. Stratford, D. L. Knol, L. M. Bouter and H. C. W. Vet. 2010. 'The COSMIN Checklist for Assessing the Methodological Quality of Studies on Measurement Properties of Health Status Measurement Instruments: An International Delphi Study'. *Quality of Life Research* 19(4):539–549.
- Moore, M. H. 1995. *Creating Public Value: Strategic Management in Government*. Cambridge: Harvard University Press.
- Moore, M. H. 2007. 'Recognising Public Value: The Challenge of Measuring Performance in Government'. In J. Wana (ed.), *A Passion for Policy* (pp. 91–116). Canberra: ANU E-Press.
- Moore, M. H. 2013. *Recognizing Public Value*. Cambridge: Harvard University Press.
- Mulrow, C. D. 1994. 'Rationale for Systematic Reviews'. *British Medical Journal* 309(6954):597–599.
- Nabatchi, T. 2011. 'Four frames for understanding public values in administration and governance', Paper presented at *Public Management Research Conference*, New York: Syracuse University, June, Available from http://publicvaluesconference.weebly.com/uploads/8/8/9/4/8894371/nabatchi_-_four_frames_for_understanding_public_values_in_administration_and_governance.pdf [Accessed 1 March 2016].
- O'Flynn, J. 2007. 'From New Public Management to Public Value: Paradigmatic Change and Managerial Implications'. *Australian Journal of Public Administration* 66(3):353–366.
- Parasuraman, A., L. L. Berry and V. A. Zeithaml. 1991. 'Refinement and Reassessment of the SERVQUAL Scale'. *Journal of Retailing* 67(4):420–450.
- Parasuraman, A., V. A. Zeithaml and L. L. Berry. 1988. 'SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality'. *Journal of Retailing* 64(1):12–40.
- Petticrew, M. and H. Roberts. 2008. *Systematic Reviews in the Social Sciences: A Practical Guide*. Oxford: Blackwell Publishing.
- Prebble, M. 2012. 'Public Value and the Ideal State: Rescuing Public Value from Ambiguity'. *Australian Journal of Public Administration* 71(4):392–402.
- Productivity Commission Australia. 2016. Report on Government Services. Available from <http://www.pc.gov.au/research/ongoing/report-on-government-services/2015/justice/police-services/rogs-2015-volume6-chapter6.pdf> [Accessed 1 March 2016].
- Rhodes, R. A. W. and J. Wana. 2007. 'The Limits to Public Value, or Rescuing Responsible Government from the Platonic Guardians'. *Australian Journal of Public Administration* 66(4):406–421.
- Rutgers, M. R. 2015. 'As Good as It Gets? On the Meaning of Public Value in the Study of Policy and Management'. *American Review of Public Administration* 45(1):29–45.
- Shea, B. J., J. M. Grimshaw, G. A. Wells, M. Boers, N. Andersson, C. Hamel, A. C. Porter, P. Tugwell, D. Moher and L. M. Bouter. 2007. 'Development of AMSTAR: A Measurement Tool to Assess the Methodological Quality of Systematic Reviews'. *BMC Medical Research Methodology* 7(1):10–17.
- Spano, A. 2014. 'How Do We Measure Public Value? From Theory to Practice'. In J. Guthrie, G. Marcon, S. Russo and F. Farneti (eds.), *Public Value Management, Measurement and Reporting* (pp. 353–373). Bingley: Emerald.
- Stoker, G. 2006. 'Public Value Management A New Narrative for Networked Governance?' *The American Review of Public Administration* 36(1):41–57.
- Talbot, C. 2009. 'Public Value: The Next 'Big Thing' in Public Management?' *International Journal of Public Administration* 32(3–4):167–170.

- Talbot, C. and J. Wiggan. 2010. 'The Public Value of the National Audit Office'. *International Journal of Public Sector Management* 23(1):54–70.
- Thomas, J. and A. Harden. 2008. 'Methods for the Thematic Synthesis of Qualitative Research in Systematic Reviews'. *BMC Medical Research Methodology* 8(1):45–55.
- Uman, L. S. 2011. 'Systematic Reviews and Meta-Analyses'. *Journal of the Canadian Academy of Child and Adolescent Psychiatry* 20(1):57–59.
- Van der Wal, Z., T. Nabatchi and G. de Graaf. 2015. 'From Galaxies to Universe: A Cross-Disciplinary Review and Analysis of Public Values Publications from 1969 to 2012'. *American Review of Public Administration* 45(1): 13–28.
- Verhagen, A. P., H. C. W. de Vet, R. A. de Bie, A. G. H. Kessels, M. Boers, L. M. Bouter and P. G. Knipschild. 1998. 'The Delphi List: A Criteria List for Quality Assessment of Randomized Clinical Trials for Conducting Systematic Reviews Developed by Delphi Consensus'. *Journal of Clinical Epidemiology* 51(12):1235–1241.
- Whiting, P., A. W. Rutjes, J. B. Reitsma, P. M. Bossuyt and J. Kleijnen. 2003. 'The Development of QUADAS: A Tool for the Quality Assessment of Studies of Diagnostic Accuracy Included in Systematic Reviews'. *BMC Medical Research Methodology* 3(1):25–38.
- Williams, I. and H. Shearer. 2011. 'Appraising Public Value: Past, Present and Futures'. *Public Administration* 89(4):1367–1384.