



Business School

12th International Conference on Tax Administration
31 March & 1 April 2016, Sydney, AUSTRALIA

Never Stand Still

Business School

Taxation and Business Law

Invisible taxation: Fantasy or just good service design?

Jo'Anne Langham

UQ Business School, University of Queensland
(PhD Student)

Neil Paulsen

Associate Professor in Leadership and Management
UQ Business School, University of Queensland

Invisible taxation: Fantasy or just good service design?

Jo' Anne Langham, Neil Paulsen

Abstract

This paper introduces a new concept for design and evaluation of public services, in particular taxation. The approach is novel as it will draw from multiple domains to construct and measure Administrative Effectiveness: an alternative to traditional "service" quality.

The research explores the commonalities of service delivery in public administration with that of the private sector. Exceptional service places the customer at the centre of the experience, achieving a balance between the needs of the organisation and those of the customer. In traditional service, a relationship occurs between the customer who is the recipient of a product, assistance or guidance and the organisation selling the service. Client service may involve more than one interaction, utilising multiple channels and touchpoints. A good client experience is created when the customer achieves the desired outcome with a high level of satisfaction and quality. Service quality is achieved through good service design.

The "service" concept is not easily translated for public administration, especially the tax system. Public administration requires a high level of cooperation and willingness from citizens to participate. Citizens derive little direct reward and the relationship is usually driven by duty or obligation. Revenue collection authorities ensure that taxes are paid according to the law. Taxpayers abide by the rules and contribute the correct amount of tax. However, the ease with which taxpayers calculate and pay taxes impacts both the efficiency of the revenue collection as well as their willingness to comply. Intentional compliance is more cost effective and efficient for the revenue authority as well as the government. Less revenue is expended on a system that is predominantly voluntary and requires minimal enforcement.

However, taxation law is byzantine with a myriad of interpretations. Such complexity often benefits a litigious few whilst hindering the majority of law-abiding taxpayers. The challenge is to design the system to hide this complexity from the average citizen so they can comply without conscious thought, whilst preventing those who wish to evade from taking advantage of gaps in the law. Design of good service delivery for tax administration cannot be measured with customer satisfaction and increased revenue alone. Few taxpayers are pleased to pay their taxes. Instead new dimensions must be developed that characterise administrative effectiveness and high quality service delivery in the public service.

The paper will present a review of the existing literature and draw together the dimensions of Administrative Effectiveness. The synthesis of the existing literature will enable the development of a new theoretical model to design and evaluate administrative effectiveness.

1. Introduction

This paper outlines the theoretical development of a new model for Administrative Effectiveness. A model for the design and evaluation of effective service delivery from a public sector organisation at the client experience level does not exist. Many public sector organisations are currently directing effort at understanding the needs of their clients and turning these insights into possible system improvements. However, little work has been done to incorporate the lessons from other domains that are advanced in their understanding and application of purposeful client experience design and delivery of contemporary client service.

I seek to construct a new model of evaluation for a modern public administration by combining the existing knowledge from services management, new public governance, engineering and psychology. Tax compliance is traditionally reliant on enforcement, with little attention paid to the reciprocal relationship of goodwill and client service for the community. Taxation administration is a core component of public administration and is often one of the largest public organisations in any country. Tax is also exceptionally complex in its nature, with more inherent obstacles to minimise non-compliance or evasion. Tax administration in Australia is multifaceted including not only revenue collection but also the management and regulation of superannuation and associated services such as the Australian Business Register (ABR).

In developing the framework presented here, we explore areas such as human factors, safety, universal design and error prevention. These seemingly unconnected areas provide insights into alternative methods that could be applied to greatly improve the facilitation of compliance in public administration.

2. The context – the Australian Taxation system

Extensive research has focused on the determinants of taxpayer compliance: predominantly taxpayer motivation. This research is founded on the presumption that compliance can be increased by influencing an individual's willingness to pay tax, in most cases by deterrence (Beccaria & Voltaire, 1801; Jacobs, 2010). Many factors motivate taxpayers to comply such as morality, ethics, opportunity, risk appetite, sense of duty, social norms, identity, attitudes and intention. Kirchler (2007) provides a detailed examination of these influences. Whilst acknowledging that citizen compliance is not solely driven by fear, taxation agencies continue to employ deterrence strategies such as audits, penalties and sanctions for maintaining high levels of compliance from the community. Audits as the primary compliance strategy are unsustainable. Audits do not treat the direct issue and have minimal flow-on effect to influence the compliance of others (Australian Taxation Office, 2014a). It is impossible to ensure sufficient audits are conducted to guarantee revenue and the return on investment in audit is poor (for every dollar spent a maximum of \$6 dollars is collected). Additionally, revenue agencies are now forced to cope with shrinking

workforces and reduced government funding (Commonwealth of Australia, 2014; Hutchens, 2014).

The existing high levels of willing participation in countries such as Australia¹ are due to trust and respect, relying heavily on the goodwill of citizens and their cooperation with the tax agency. High levels of intentional compliance is not maintained by force, but instead by community willingness and cooperation (Kirchler, Hoelzl, & Wahl, 2008). Therefore, it is unlikely that increasing the frequency or scope of audits will improve compliance. The recognition of the importance of willing participation means that many contemporary taxation administrations are now attempting to positively influence taxpayer intention through greater engagement and communication with the community.

However, the view that citizen decision making or taxpaying choices can and should be influenced by the government, assumes *taxpayers ordinarily choose to evade* rather than pay their taxes. This assumption is incorrect. In a great deal of tax compliance literature the majority of taxpayers in many countries *already* choose to comply (Schneider, Buehn, & Montenegro, 2010). A small minority attempt to avoid their obligations and use various means to evade taxes and challenge the system. These ‘seriously non-compliant’² taxpayers are those who are not easily manipulated and intend to evade. Robust enforcement approaches are required to corral these individuals. However, the majority of taxpayers exist between the extremes of high competence and deliberate evasion. These taxpayers are either influenced by social norms and opportunity to leverage loopholes in the law, or in the majority of cases, try to do the right thing but fail.

This large slice of the taxpaying population is where major improvements in tax compliance can be made. Complexity of law, ambiguity of process, poorly designed system or product interfaces all contribute to non-compliance. The tax system itself is what often leads well-meaning taxpayers to make unforced or accidental errors. The robustness and quality of the administrative design is an important but overlooked component of the compliance dilemma. Significant evidence shows that tax compliance is a complicated and costly endeavour for many taxpayers, particularly those in small businesses (Hasseldine et al., 2012). This burden is exacerbated by the manner with which taxpayers must interact with the bureaucracy to meet their obligations.

The tax system has been stringently designed to deal with the litigious few, rather than to provide simplicity and clarity for the majority of compliant taxpayers. The law is constructed to prevent illegitimate use of legal gaps or loopholes. Drafters of taxation law take great care to ensure that fewer paths are available for those who choose to act in ways that are contrary to the spirit or intention of the legislation. However, the consequence of this concentrated effort means the client or service

¹ Non-compliance is assumed to be as little as 3-4% in Australia (Australian Taxation Office, 2009)

² Tax Office terminology used to describe deliberately non-compliant taxpayers.

experience is a secondary consideration to the protection of revenue. Precision and accuracy of law design has until recently, been considered paramount. The focus on legal correctness has inadvertently led to high levels of complexity and poor user experience. Complexity of law is usually born by the taxpayer as high compliance costs (McKerchar, 2003a, 2003b). Very few people have the resources, expertise or knowledge to correctly interpret and apply the law to their circumstances. Obviously those taxpayers with greater resources are more able to cope with legal complexity.

However, it is not only the law that makes the taxpayer experience so difficult. Evidence indicates that the tools, education materials and the interfaces to the tax systems are difficult to use, hard to interpret and may actually impede compliance. This consequence is acknowledged in recent community engagement from the Australian Treasury through its white paper *Re:think* (The Australian Government the Treasury, 2015). The consultation process between the government and the community seeks to gain insights to resolve the issues around both legal and administrative complexity. The whitepaper concedes that a functioning tax system should be equitable, provide certainty and simplicity in participation (The Australian Government the Treasury, 2015, p. 14). There are 795 publicly available submissions to the review. Many of these are focussed on the legislation and policy related to tax, rather than the administrative system and the complexity of the process. Complexity should be addressed at the source if possible. Obviously, simplifying the law will undoubtedly make the tax system easier to administer. However, there have been many attempts to review the tax system and little success in actioning the recommendations of the various review processes.

Regardless of the difficulty of law, even theoretically simple tasks, like paying a known amount of tax can be made complex through poorly designed administration. The service provided by the organisation must still be designed understanding the client-citizen experience and ensuring that obvious obstacles to compliance are removed. Online services such as phone banking have been available since the mid 1970's with personal computer access since the 1990's (Chou & Chou, 2000). These services have evolved significantly with improved accessibility and ease of use since the original design. Businesses with websites grew to provide payment services built on the conceptual template provided by the banks. Payment through online interfaces are now commonplace. However, these experiences can still be clumsy and cause error.

The administration of these types of conventional tools in government organisations such as the Tax Office often demonstrates an unnecessary amount of complexity. For example, when a taxpayer attempts to pay their taxes through an online system they may accidentally pay the wrong account (for example: Goods and Service Tax [GST] rather than the Pay as You Go [PAYG] account)³. The system user interface does not make it clear to the taxpayer that multiple accounts exist or that a mistake could be made. Such an accident is compounded as the system does not automatically inform

³ information as at April 2014.(Australian Taxation Office, 2014b)

the taxpayer that funds have been credited to this account, as a Statement of Account (SOA) is not issued unless requested by the taxpayer. The taxpayer may continue to incur interest and penalties until a certain threshold is reached (approximately \$12,000) and only then will a statement be despatched. By this time, the debt is critical and the taxpayer may be unable to pay (Australian Taxation Office, 2014b). The expected and common behaviours or patterns that take place in commercial digital products are not necessarily replicated in public service systems. The warning mechanisms that alert users to errors are not always present in these systems.

Complexity or further obstacles to compliance can be introduced to the administration in other ways. The use of tax jargon is also a pervasive issue, often obscuring simple tasks. A lack of understanding of obligations and the related impacts can be seen in the management of General Interest Charge (GIC). GIC is a substantial component of collectable tax debt. Most taxpayers do not know that GIC is a daily compounding interest rate (Millward Brown, 2014a).

The intention of GIC is to encourage taxpayers to pay their debt quickly. However, due to complexity it means that many do not understand the nature of the sanction, thus negating the intention of this charge. In 2014, the Tax Office undertook research with taxpayers in debt to determine their attitudes and behaviours relating to debt and to understand the reasons for the increasing rate of tax debt. Only 61% of Small and Medium Enterprises and Micro businesses currently in debt were aware that the payable interest was calculated on a daily compounding basis and only 29% knew that the current interest rate was between 9-10% (9.6%) (Millward Brown, 2014a). GIC accounts for 15% of collectable debt for clients with outstanding lodgement and 20% for clients that have two to five outstanding lodgements. This increases to 30% for those with five or more outstanding lodgements. The results clearly indicate that the average taxpayer does not comprehend GIC. The systems with which taxpayers interact and the letters they receive do not make it easy to understand the charge or what it means for their overall tax position.

Even though taxpayers may want to comply, many struggle to achieve this goal. In Australia, each year many complaints are made to the Tax Office as well as external scrutineers such as the Inspector-General of Taxation and parliamentary committees. These complaints relate predominantly to inadvertent non-compliance. In 2013–14 the Taxation Ombudsman⁴ received 1,369 complaints about the Tax Office (Commonwealth Ombudsman, 2014). During the same period the Inspector-General dealt with issues relating to the income tax refund integrity program, data matching, poor engagement and the inappropriate application of penalties for self-assessment reviews (Inspector-General of Taxation, 2014a, 2014b).

⁴ In 2014 the complaints role of the taxation ombudsman has been transferred combined to the Inspector-General of Taxation. Now all community complaints (not handled by the ATO) are handled by the Inspector-General of Taxation. Categories of recent complaint figures from the general community are not available since the integration of roles.

Debt management is considered a significant problem for the tax administration. In 2005, The Inspector-General conducted an inquiry into the Tax Office's debt collection process which revealed a number of issues of concern, many of these were due to ambiguity and difficulty relating to rigid processes and procedures (Inspector-General of Taxation, 2005). The Inspector-General of Taxation conducted another review in 2014 as Australia's tax debt continued to grow to more than \$20 billion in 2013-14. Key recommendations included creating systems that made it easier for taxpayers to manage their debts, capacity to pay tools, remission of interest to encourage future payment, and processes to correct data mismatches with garnishee notices.

A similar issue is the perception of a lack of fairness in the Tax Office's disputes handling. Recent research (Millward Brown, 2014b) reveals general taxpayer dissatisfaction with the Tax Office processes for correcting mistakes or resolving issues that have occurred due to administrative failure. Lack of procedural justice is one of the strongest factors of taxpayers' perception that they are being treated unfairly⁵. The government has launched an inquiry into the disputes handling process due to claims of⁶: delays; undue suspicion; no transparency; lack of clarity in guidelines; disrespect by tax officers; complexity and lack of process uniformity; and lack of natural justice (Bevacqua, 2015; *Inquiry into Tax Disputes*, 2014).

These few examples show the basic effect of complexity and the complicated administrative systems on taxpayer behaviour. The impact of the 'tax system' on the performance of tax compliance behaviour is considerable. The introduction of system inhibitors or obstacles to compliance, such as poor usability, directly influence not only the decision making process but also the ability for the taxpayer to correctly comply. The actual design of the administration or the tax system must be considered when determining causes or influences on compliance behaviour. The effect of the design of the administration, including the tools, forms and processes the taxpayer must use, are persistent and unavoidable. The design of the administration mechanisms impact all tax compliance behaviours from the simplest, such as reading a letter from the Tax Office, to the most complex, which may involve interpretation of the law.

The Australian government is attempting to consider effectiveness of taxation administration in terms of impact at the behavioural level through a focus on complexity. The Treasury's taxation review "Re:think" poses the option for measuring or developing metrics for administrative complexity (The Australian Government the Treasury, 2015, p. 171). Understanding and measuring complexity within the tax administration, demonstrates a shift in thinking away from organisational performance measurement towards understanding and facilitating compliance for taxpayers. Incorporating a measure of complexity is a noteworthy change in direction. However, complexity is only one aspect of administrative

⁵ Only 41% of those in dispute in with the ATO (2014) thought that the process was fair (Millward Brown, 2014b).

⁶ Taken from various submissions to the Standing Committee on Tax and Revenue – Inquiry into Tax Disputes

effectiveness. Compliance failure is caused not only by administrative complexity, but also by misunderstandings, fragmentation of experience, delays and other experience obstacles. Therefore the metrics must include broader aspects of the service delivery.

Further, the client experience must be considered as a whole. For example, a task such as reporting income tax which can be completed using an online tool such as myTax, should be designed around the entirety of the experience for the taxpaying individual. The process of reporting one's income tax does not begin at the point when the application is launched and fields are completed. Reporting income tax begins at the commencement of the financial year, where records are kept and advice is taken on how to manage these records, what to keep and what to ignore. Reporting and managing the evidence for reporting income tax is a year-long endeavour. There may also be other concurrent threads of activities in this overall experience, such as calls to a helpdesk, or advice from an accountant. Each of these elements, which may involve other artefacts or tools are part of the experience of "reporting income tax". The public service is challenged in determining the boundaries of such an experience. Where is the conclusion of the experience for reporting income tax? Does it end at an online receipt or confirmation from the online tool, at the point of a Notice of Assessment (NOA) or is there even an end point? Is income tax reporting an ongoing process? This would mean that the usual linear flow of a client experience must be reconsidered. These complex blueprinting notions scarcely exist in the public service design context.

Additionally, administrative organisations or bureaucracies such as the Tax Office are not structured to enable the client experience to be at the forefront of the design process. The larger the organisation becomes, the more artificial the internal structures to manage the delivery and operation of the organisation. Authority and accountability are aligned to the delivery of isolated components of the administrative design. Individual areas are rewarded for the delivery of their component not the improvement to the overall client experience. Less emphasis is placed upon ensuring an holistic and seamless experience for the taxpayer. Although organisations such as the Tax Office are now including these considerations in their future planning processes (Australian Taxation Office, 2013, 2015; Jordan, 2014), valuing and measuring the client experience is still a difficult and incomplete endeavour. Large bureaucracies require good governance systems to undertake such significant strategic changes such as the move towards a client centric focus. The governance metrics must be tangible, measurable and defensible. The measures must also enable government to connect performance, delivery and outcomes. Measurement of performance drives the organisational behaviour. However, there is a significant gap in connecting measures of organisational performance with improving the client experience. A true measure of administrative effectiveness is required.

3. Administrative Effectiveness in taxation

It is the responsibility of the government and the public service to create systems that facilitate client-citizen compliance. The primary function of public administration such as the Tax Office, is ‘service delivery’. However, taxpayer obligations are mired in convoluted administration. Poor design minimises the influence of taxpayer purposeful intention to comply. Strategies directed at influencing intention become redundant when the system itself does not support and enable the desired taxpayer behaviour. Good administrative design is instrumental in the successful performance of activities and interactions required by citizens to meet their obligations. It is not the role of public administration to command obedience and as evidence suggests, nor is this approach successful (Hoelzl, Kirchler, & Wahl, 2008).

Previous research by the lead author (Langham, 2012) exploring compliance behaviour, showed that taxpayer intention to comply with their obligations was moderated by the effect of Administrative Effectiveness (AE) (performance control). The research utilised the Theory of Planned Behaviour (Ajzen, 1991) as a basic model to determine if tax compliance behaviour could be predicted. The research demonstrated that even those completely dedicated to correctly reporting their tax, could become non-compliant through unforced errors. The administration was shown to significantly affect the overall compliance behaviour of the participants. This initial exploration indicated the need for further investigation of ‘Administrative Effectiveness’: its composition and methods of measurement and evaluation.

“Administrative Effectiveness” in this context is different to the definition provided in the majority of the literature relating to overall organisational performance and the ability to maintain and protect the revenue. No measures relating to this new variable exist. The following review of the literature will outline the relevant connections to construct a theoretical model for “Administrative Effectiveness”, an alternative measure for **service quality** in public administration.

The successful delivery of tax administration is customarily measured in terms of revenue collected in relation to the amount of funding required to achieve such an outcome (Sanford, 1994). Alternatively, effectiveness may be measured in terms of the revenue leakage or ‘Tax Gap’ (Blanthorne & Yance, 2006; Mohamed, 2012) or according to a number of benchmarks that relate to structure and process (Gallagher, 2005). These measures are all macro performance measures of tax administration. However, these measures are inappropriate at the group or individual compliance behaviour level. Using the Principle of Compatibility (Fishbein & Ajzen, 2010, p. 44) all behavioural criteria must be measured at the same level of generality or specificity. Therefore, AE of the tax system in relation to the performance of a taxpayer’s behaviour must also be measured at the behaviour or task level. In the case of the tax system and the Tax office the measure should relate to how well the organisation successfully facilitates compliance at the taxpayer behaviour performance level.

AE relates explicitly to the performance of the behaviour being examined in the client-citizen experience. For example, AE can relate to the tools, system and other supporting mechanisms that enable the individual to pay the tax required for the quarterly business activity statement (BAS). AE also encompasses the entirety of the tax compliance experience which may require more than one tool, or activity to interact with the system. In the case of making a tax payment, the behaviour would comprise incurring the liability, receiving the notice for payment, interacting with the system to make the payment and receiving acknowledgement of the receipt of the payment.

The definition of ‘effectiveness’ in this context is based on the program logic model⁷ (Funnell, 1997). The series of interactions that occur with products or services at a behaviour level that facilitates individuals achieving the behavioural objective have been known under a variety of terms such as the client experience (Niemiowski & Wearing, 2003), customer experience (OECD, 2014), citizen experience (OECD, 2014) or user experience (Mauldin, 2013; Qiu, 2014). Given the tax context, where the individual or business interacting with tax administration may be both a taxpayer and/or a client, I will term these interactions the *client-citizen experience*. This paper focuses on how researchers have understood and evaluated the client-citizen experience at the level of behaviour as well as how this information has been used to create new improved services in a variety of domains. Existing knowledge on designing and measuring the facilitation of client-citizen behaviour will assist in constructing and measuring AE for the compliance behaviour model.

Few studies objectively examine tax agencies or comparable large public organisations to assess their claim of effectiveness in terms of how well they deliver specific services to their constituents. This manner of assessment is how commercial organisations are intrinsically evaluated. Service delivery of the client-citizen experience and customer service is what ultimately determines the success of commercial enterprises. Although some would claim that the new public management approach has more natural accountability (Kaboolian, 1998), public organisations are moreover precluded from the real world test of the unforgiving market forces that consume sub-standard businesses.

The public sector cannot be assessed by tangible measures such as profits. Profit is the outcome of efficiency, quality and various factors contributing to the performance of the organisation. No equivalent outcome measure is available in the public sector. Public services are predominantly social exchanges with the community, where citizens participate to both contribute and receive benefits. The closest possible equivalent value is compliance with the law. Revenue collected by the tax authority is too basic a measure, as it does not account for the broader service ecology. Revenue is closer to a measure of efficiency, rather than effectiveness as the manner in which it is collected can vary greatly and may require coercion and extreme effort to collect. In

⁷ **Program logic** is a theory about the causal linkages among the various components of a program: resources; activities, outputs; and short and long term outcomes (Funnell, 1997)

an effective system, the correct amount of tax revenue should be easily collected and with minimum adverse impact on the taxpayers.

Regardless, improved service does not necessarily equate with increased revenue. In commercial enterprise we can observe a relationship between improved service, higher sales and increased profits (Garvin, 1983; Heskett, Jones, Loveman, Sasser, & Schlesinger, 1994). The entrepreneurial spirit that drives innovation, organisational and service improvements is absent from most public service organisations such as the Tax Office. Improved service in the tax system means that taxpayers will be more likely to pay the correct amount of tax and in some instances, even less tax (McKerchar, 2002).

Alternatively, a result could be more “indirect revenue” as opposed to “direct revenue”. Direct revenue is the cash collection or liabilities raised through audit procedures. These are tangible, easily defined measures. Measurement of direct revenue is currently the primary method of determining tax agency performance. However, the increase in indirect revenue, which is created through improvements in automatic voluntary payments by taxpayers, is difficult to determine. The source of any upturn is unclear and the connection with compliance strategies problematic. Therefore, incentives for service improvement must come from other sources (Pollitt, 1993, p. 124). The current ‘Reinvention’ program for the Tax Office (Jordan, 2014) was instigated through a change in Commissioner as well as the threat of the organisation being dissected and separated due to issues such as poor service and lack of dispute resolution. In 2013, the Treasurer, Joe Hockey established the parliamentary committee to oversee tax administration, known as the Standing Committee on Tax and Revenue. The creation of an external committee to scrutinise performance was the government’s response to community sentiment that the Tax Office could not accurately evaluate its own effectiveness or provide an impartial view on whether it treated taxpayers fairly.

Although the service delivered through client experience is not directly measured, it is clear that the process, procedures and interactions that the public have with the Tax Office are distinct factors to consider in the effectiveness of the organisation. However, unless these processes are specifically pursued and evaluated, they can run counter to traditional performance measurement and evaluation of a public service organisation.

The purpose of the current research is to deconstruct the concept of the client experience in the public service and demonstrate the impact the various dimensions of administrative effectiveness have on the service outcome: compliance behaviour. The development of a practical and parsimonious model for administrative effectiveness and its role in compliance behaviour will enable public authorities to deliberately design good public administration. Such a model can then be evaluated and applied practically to improve compliance. At a whole of government level, this information can be utilised to improve services, the client-citizen experience and support improved government. The application of this knowledge will also contribute to the fairness of the tax system

and other government services. Hiding complexity and ensuring all have access to the same level of service will improve equality and the integrity of the system for all.

Future research may also utilise the variable of AE to measure environmental impacts on behaviour rather than continuing to use proxy measures which often leave large unexplained errors in the testing of behaviour models. The estimation of an individual's ability to affect a desired behaviour spans the gap that has been proclaimed in psychology between intent or self-interest and actual behaviour (Miller, 1960; Nisbett & Ross, 1980, p. 11).

The following review of the literature will focus on relevant disciplines that can provide insights into devising a core set heuristics to evaluate the design of the administration in regards to supporting and facilitating good compliance behaviour. The literature will be used to isolate possible dimensions that could be included in a model for effective administration design. This in turn will provide a path to measure and improve public administration and enable bureaucracies such as the Tax Office to improve their administrative design processes as well as evaluate and enhance their client's experience of the revenue authority's service. The development of a core set of heuristics may also provide a germane set of key performance indicators to drive the behaviour of management and staff operational delivery.

4. The literature

The primary areas of exploration are new public management; new public governance and services management. These areas have all developed during the same period and provide various insights into the delivery of both private and public services. All of these disciplines explore the relationship between the consumer and producer of a service, which are integral to the theories. However, there has been little cross-pollination of ideas and therefore extensions to the theories that can readily be utilised for the design and evaluation of effective tax administration.

In general, the successful delivery of citizen focussed service and the new design methods that are being employed by organisations have evolved from human factors and ergonomics, services marketing and service design, and utilise public participation methods such as user-centred design, co-design and co-production. There is a natural fusion occurring between the use of these methods and a desire by public administration to form a nexus with community participation and to create a new manner of public administration. This is occurring as the application of service and design theories forces public administration to question how and what they deliver and in turn how they are best organised to achieve these goals (Jordan, 2013, 2014). Governments are moving beyond individual organisations providing specialised services, to holistic client-focused delivery mechanisms. In Australia, one such example is myGov.au (and subsequently myTax). Such client-centric tools are indicators of a burgeoning shift in public administrative thinking.

Through understanding the various contributions of these diverse fields a new framework for designing and assessing the effectiveness of public administration service delivery, specifically tax service delivery, will emerge. I will then outline the key variables for consideration based on the extant research that contribute to understanding, measuring and designing taxation AE.

Public administration

Public administration is distinct from the field of public policy as it is concerned with implementing policy, not only the development of law ("Public Administration," 2008). Although resourced using civil servants who are employed through merit based processes and who are meant to be politically neutral, the decision making and implementation of policy is often politicised (Ashworth, Boyne, & Entwistle, 2010). It is well recognised that the very environment of the public service organisation is political.

Public service organisations are traditionally, bureaucratic, rigid, and durable. They have multiple layers of management with narrow spans of control and specialised functions (Golembiewski, 1969). Generally, administrative delegation is mired in legislation and detailed in minutia (Golembiewski, 1969) which restricts self-direction and control of organisational change. Characteristics that are valued in public administration include accountability, managerial control, equality, security and procedural regularity (Golembiewski, 1969; Levine, Backoff, Cahoon, & Siffin, 1975).

Measurement of effectiveness has rarely been documented or assessed with rigor at anything but the program or organisational level. Early scholars of public sector administration focussed on industrialisation and the use of performance management of individuals in the workforce (Murphy, 1995). This evolved to a more strategic approach to overall outputs – but still focussed on the measurement of ‘widgets’ in the early 1800’s (Heinrich, 2009). Keynesian economics dominated government management until the early 1970s which was overtaken by economic liberal thinking and the ‘New Right’ (Farnham & Horton, 1993). These ideologies had a profound effect on the direction of the public service. Keynesian thought was underpinned by collectivism, equality, social rights and the belief that the government was required to intervene to balance the market forces and the economy. The New Right castoff the welfare cloak of the post-war years and instead focussed on individualism, personal freedom and faith in the natural balance from the market forces (Farnham & Horton, 1993). The political direction of the government during these times lead to changes in the management and evaluation of the public administration.

Through the 1940’s to the late 1980’s there were a number of approaches to evaluating or measuring the performance of the public sector. Several of these approaches were due to reforms that were focussed on improved management and incentives based rewards (Barnard, 1938). There were a number of management reforms including the Planning, Programming, Budgeting System (PPBS),

Management by Objectives (MBO) and Zero-based Budgeting. These were all efficiency focussed or “*output*” measurement approaches used to manage and evaluate performance (Heinrich, 2002).

In the late 1980’s Total Quality Management (TQM) emerged as a challenge to the approaches of efficiency focussed measurement (Deming, 1986). TQM focussed on the quality of results or ‘*outcomes*’. This was the first apparent period when the view of customers or clients of the performance of the organisation was seen as important. However, tangible measures based on client views were seen as difficult (Heinrich, 2009). Customer satisfaction was volatile, and a reliance on a community that may not be well informed to judge the performance of the public sector, was not popular. The public was seen as having ‘ambivalent views’ and ‘unable to provide reliable feedback’ due to differing values, economic self-interest and racial distinctions (Beck, Rainey, & Traut, 1990).

Similarly in the 1980s the United Kingdom and the United States employed the Financial Management Initiative (FMI) which was based on the principles of Value For Money (VFM) focussing on ‘economy’, ‘efficiency’ and ‘effectiveness’ (Osborne, Bovaird, Martin, Tricker, & Waterston, 1995). There was a movement in the public sector towards privatisation, business planning, citizen views, independent inspectorates and closer control by a central government (Osborne, et al., 1995). Even though progress had been made, there was still a heavy reliance on the rationalistic perspective for evaluating public administration.

Public administration as a field merged with that of public management in the early 1990’s and became more focussed on the practices and training of executives and managers. It was also considered to be a sub-field of management education (Jones, 2012). This gave birth to the field of ‘managerialism’. Managerialism saw a change in education of public administrators to be focussed on more ‘business-like’ practices including the role of managers (Midwinter, 1990). There was strong critique of this shift in thinking by educators and a number expressed their views in editorials of the change both then (Davies, 1998; Midwinter, 1990) and continuing now (Jones, 2012). Although directed at measuring performance in a number of ways, there was a keen interest on the involvement of citizen views, ‘customer orientation’ and ‘putting citizens first’ (Davies, 1998).

This was a noteworthy change with the rise of the New Public Management. The New public management took the first steps towards investigating compliance behaviour from the citizen experience perspective.

New Public Management (NPM)

New public management theory suggests that public administration should model itself on the structure and management processes of the private sector. The NPM approach demonstrates a move towards ‘de-politicisation’ and ‘de-bureaucratisation’ of the administration of the public service (Hood, 1991; Maor, 1999).

NPM, which at its core embraces the philosophy of managerialism, implies that through good organisation and 'good management', the public service could be pushed to achieve and progress beyond its previously officious constraints (Gray & Jenkins, 1986). Efficiency and productivity are clear factors of success for business and so the general belief is that if applied to the public service, these approaches will create substantial improvements. Advances will be made if senior public servants become "professional managers"(Pollitt, 1993, p. 2).

One of the clear developments in NPM is an acknowledgement that performance evaluation should not only look at the outputs but also to the impacts of the services. As Osborne et. al. (1995, p. 28) explains: "*impact monitoring, involving the assessment of the achievement of the highest levels of organisational objectives, normally in terms of the impacts (outcomes) achieved on the targeted client population or organisations. It is difficult, and therefore generally occurs less frequently, though conversely it is one of the most significant forms of performance assessment.*"

The New Public Management ensures a distinct shift in focus towards client-citizen service which requires evaluation of customer satisfaction and quality of service. Understanding the satisfaction of 'customers' implies that they receive a service or a product. However, this assertion does not extend to assessing or evaluating the objective experience of the citizen receiving these services. Service satisfaction and quality measurement rely heavily on the subjective perceptions of customers. Additionally, contradictory elements to the new public management claim of customer service orientation also exist. For example, the theory does not acknowledge that the government in general has a monopoly on particular services and products. Consequently, the citizen (or customer) has no alternative to choose if the service is poor, unfair or frustrating (Farnham & Horton, 1993, p. 37).

Further paradoxes in NPM have created inconsistencies and impaired the growth and interpretation of the effectiveness of administration and the role that it plays in citizen compliance behaviour. A focus on productivity improvements for an organisation often, and sometimes contradictorily, means staffing or resourcing reductions. Whilst this approach may be appropriate in some areas of government that have become conspicuously inflated, when applied broadly it also means a reduction in skilled and technical specialisation as well as in areas that rely on a human interface, such as shopfronts, call centres or helpdesks. Areas in service businesses that have a positive influence and impact on customer satisfaction are usually those that involve a face-to-face interaction directly with the client (Gabbott & Hogg, 1998, p. 68). The human interaction element is seen as a vital element in satisfaction (Quelch & Ash, 1981). Interpersonal contact is also the best opportunity to provide tailored, individualised service. Therefore, reductions in these areas in public administration can cause systemic issues in the customer or client experience.

Public administration also differs from many commercial organisations in that it has long term interactions with their citizen-clients and for this to be of value the

relationship requires trust and develops over time. These relationships are an essential element of the effectiveness of public service. Pollitt (1993, p. 128), states that the relationships are not necessarily there to “satisfy” the citizen-client; they are often there to shape, change or encourage certain behaviours. This aspect is applicable with regards to encouraging taxpayers to meet their obligations. It is not a requirement of good tax administration for taxpayers to feel ‘satisfied’ with paying taxes. However, the manner in which they achieve the goal of compliance should be considered a service delivery. Deficient service from the Tax Office creates a sense of unfairness for taxpayers. Poor procedural justice leads to dissatisfaction, disengagement and in extreme cases, defiance (Braithwaite, 2009). Therefore, poor client service can actually initiate non-compliance.

Although recognising the need for service evaluation, no method other than quantitative “customer satisfaction” approaches have been developed through new public management. The contemporary tax agency assesses customer satisfaction with the organisation at a broad program level. The Tax Office has been conducting perception and satisfaction surveys in different forms for a number of years. From 1996 to 2013 the Tax Office assessed the general interactions with micro-business⁸ and individuals through the Community and Micro-enterprises Perception Survey, small businesses⁹ through the Small and Medium Enterprises Perception Survey and tax agents through the Professionalism Survey. Currently, the Tax Office conducts a combined customer satisfaction survey covering all market segments known as the Single Corporate Perceptions Survey¹⁰. These surveys evaluate general measures of performance in regards to channel delivery such as digital service (website, portals and apps), shopfront service and call centres. These measures are used to evaluate the performance of the Tax Office against service commitment standards.

Another exploration of service quality for taxation in the new public management approach borrows from the services management literature. These studies utilise well known scales such as SERVQUAL (Kakouris & Meliou, 2011). No set benchmark for quality service or an objective measure for service currently exists (Parasuraman, Zeithaml, & Berry, 1988). However, the service management literature uses proxy measures to indicate service quality. SERVQUAL measures the difference between customer expectations and whether these expectations have been met and serves as a measure of good service (Alexandria & Adrienne, 2001; Parasuraman, et al., 1988). Parasuraman, Zeithaml and Berry developed the SERVQUAL scale to provide a useful measure of service quality. Their research presumes the foundations of measurement based on customer or consumer attitude. They discuss the additional components of satisfactions, expectations and perceptions. All of these variables are similar to attitude.

⁸ Micro-businesses are defined as businesses with a Total Business Income (TBI) of \$2 million or less

⁹ Small and Medium Enterprises are defined as businesses with a TBI of \$2 to \$250 million

¹⁰ From 2014

Few examples exist of how SERVQUAL has been applied to the service provided by a tax administration. Pinho, Macedo and Monteiro (2007) utilised SERVQUAL to assess tax accountants' view of online customer service for the Portuguese tax department. The researchers' aim was to identify issues and assess the website so that it "*conforms to rigorous quality standards in order to facilitate user compliance with fiscal obligations*". The research question was:

To what extent do online services provided by the Department of Taxation web site meet quality perceptions?

Five core constructs were derived from previous work in service quality (Parasuraman, Zeithaml, & Berry, 1985), and specifically e-service quality (Yang & Jun, 2008), for the measurement of satisfaction and quality of the online service. The original constructs were: *tangibles; reliability; responsiveness; assurance* and *empathy*. A further construct of *convenience* was also tested. The results demonstrated a strong relationship between satisfaction and empathy, tangibility and convenience. However, the authors acknowledged a number of limitations including generalizability of the findings. Additionally, they noted that *satisfaction* is constructed by both cognitive and affective elements; *service quality* is predominantly a cognitive construct. The authors claimed that *satisfaction* and *quality* are distinct, but over time the two constructs converge. This proposition has also been supported by others (Dabholkar, 1995; Iacobucci, Ostrom, & Grayson, 1995). However controversially, de Ruyter, Bloemer and Peeters (1997) demonstrate instead that service quality is instead an antecedent to service satisfaction, and that these two constructs are in fact distinct.

Despite the extensive research supporting the service quality field, an important ontological distinction exists between the quality and satisfaction constructs discussed here and the principal construct of discussion: AE. Service quality and satisfaction are internalised 'perceptions' of quality and satisfaction, which can only be known by the customer-consumer. However, AE, relating to the actual quality and performance of the service is an externalised concept and may be independent of the perceptions of the customer-consumer.

The existing scales from SERVQUAL cannot be used in the context of the Compliance Behaviour Model for several reasons. Humans have cognitive biases which distort reasoning. Biases such as cognitive fluency (Unkelbach, 2006) will lead people to trust or believe in experiences due to frequency and fluency or the framing effect (Sunstein & Thaler, 2003) which enable people to draw conclusions based on the way information is presented. Such blind-spots in human thinking mean that self-judgement of an absolute or intrinsic value of quality of service will be shaped by a number of factors that may create false conclusions. Secondly, objective self-reporting of behaviour is well known to be unreliable (Hessing, Elffers, & Weigel, 1988). Self-reported behaviour is affected by issues such as common rater bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003), social desirability, imperfect attention (Johnson, 2010, pp. 87-105) and poor memory (Loftus & Palmer, 1974).

Objective assessment of performance cannot be self-reported with any level of accuracy, instead this must be observed by a third party (Nielsen, 2001).

Service quality dimensions of *tangibles; reliability; responsiveness; assurance* and *empathy* are possible external measures and need to be reconceived with appropriate operationalisation. The services literature is able to provide useful insights for the conceptualisation of an AE construct however it is incomplete as an holistic measure.

Service management, design and the new public management

The increasing focus on ‘service’ delivery of the new public management occurred in parallel and isolation to the development of the services management literature (Osborne, Radnor, & Nasi, 2013). The services management literature developed as an extension to marketing theory (Gronroos, 1978). Although it is debateable whether commercial and public services organisations are similar in service provision, analogous conclusions can be drawn about effectiveness and client experiences.

Many definitions for ‘service’ are possible, although the general consensus is that services involve the exchange of intangible resources, relationships and co-creation (Vargo & Lusch, 2004). Formal characteristics are provided by Sampson and Froehle (2006) as part of the Unified Services Theory (UST). These characteristics of service are: *heterogeneity, perishability, simultaneity, intangibility and co-production*. (Sampson & Froehle, 2006). The UST was the emerging all-encompassing services theory that incorporated key factors from those such as: the service process matrix (Schmenner, 1986); a taxonomy of service (Wemmerlöv, 1990); the service process/service package matrix (Kellogg & Chase, 1995); molecular model (Shostack, 1982); service blueprint model (Shostack, 1984); Service-Provider-Customer (SPC) (Mo, Xu, & Wang, 2008); Structured Analysis and Design Technique (SADT) (Congram & Epelman, 1995); and dynamic event process chain (EPC) (Mo, et al., 2008).

Osborne, Radnor and Nasi (2006, 2009; Osborne, et al., 2013) have provided some direction towards a reconciliation between the new public management and the services management literatures. They have termed this work New Public Governance (Osborne, 2006). The previous approaches to public service management have been based on the industrial or manufacturing paradigms. Such approaches treat the services provided by the government as individual products which can be consumed and are distinct. This view has been noted as a “product-dominant fatal flaw” in the public management theory (Osborne, et al., 2013). New Public Governance (NPG) recognises the facilities for what they are: public *services*. Osborne (2009, p. 6) describes the thinking for NPG as:

“both a pluralist state, where multiple processes inform the policy-making system, and a plural state, where multiple inter-dependent actors contribute to the delivery of public services.”

Public services in this new domain are considered as complex, multifaceted interactions not only between the consumers of the services (citizens) and the public

administration, but also amongst the numerous government organisations that contribute to the citizen-client experience (Osborne, et al., 2013). The recognition of the holistic experience is a considerable step forward in the theory.

Osborne et al (2013, p. 139) conclude that a focus on service delivery or a “service-dominant” rather than “product-dominant” approach, means that the client-citizen experience should not be considered as peripheral in the management, design, or evaluation of the service; instead it is central to successful public service delivery.

“The performance of a public service is not (solely) about its effectual design in relation to its purpose but is, at least, equally about the subjective experience of that service by its users.”

Consequently, ‘coproduction’ of services is intrinsic to the new public governance approach. However, Osborne et al (2013) criticise the current use of engaging users in the coproduction of services as it has been introduced through product-dominant or manufacturing techniques. The outcome of this misapplication has led to superficial understanding of the user’s knowledge and often leads to erroneous conclusions. Product-dominant designer expectations that the users’ knowledge of current activities and tasks can be used to improve and anticipate future requirements, are unrealistic. The services literature is more advanced in co-production and has incorporated an understanding of how the users of the service, as they interact with the service, respond and shape future experiences (Gabbott & Hogg, 1998, p. 89). The services literature’s incorporation of symbolic interactionism as a theoretical basis is absent from the new public management research. New public governance attempts to advance the thinking in this regard by proposing the coproduction process has a greater focus on the tacit or “sticky” knowledge of users and that they are central to the design process (von Hippel, 1994).

“By taking a public service-dominant approach, coproduction becomes an inalienable component of public services delivery that places the experiences and knowledge of the service user at the heart of effective public service design and delivery.” (Osborne, et al., 2013, p. 146)

Researchers have proposed alternatives to the new public governance such as public value management or collaborative government; outcomes management or alternatively public value pragmatism (Alford & Hughes, 2008). These approaches emphasise cooperative, networked arrangements and flexibility in applying frameworks to improve public services. Alford and Hughes (2008) argue that the next step in public management and administration is about pragmatism, flexibility and a focus on problem solving for public value.

A greater depth of understanding now exists in recent new public management literature in regard to the citizen-client relationship with the government. Several researchers have claimed that the services doctrine devalues the ‘citizen’ role and that citizens are viewed as passive agents in the relationship (Alford, 2002; Frederickson, 1997). Citizens are not customers per se, as this approach implies a one way flow of

services, but instead citizens act within a form of social exchange – where they receive goods and services from the government and in response willingly comply with certain obligations (Alford, 2002). This view is prevalent in the literature and some associate it with symbolic interactionism (Blumer, 1969; Gabbott & Hogg, 1998, p. 82; Goldkuhl & Lind, 2009; Stryker, 1980). The concept of an evolving relationship between the government or tax authority and the citizen, based on changing signals and interactions has been explored in the tax context, although not in relation to service encounters (Rothengatter, 2005).

The interaction between the tax officer and the taxpayer has a significant influence on compliance behaviour. The stereotyped roles of “cat and mouse” are considered to be a product of the continuous interaction and feedback and in turn these roles further shape the adversary nature of the relationship (Rothengatter, 2005). The application of symbolic interactionism must be considered within the public service provision context and how the relationship can develop to influence the service experience. The interactive two-way flow of information as well as co-operation amongst parties must be included as part of any model that seeks to incorporate co-production and the dynamic effect that each participant brings to the service experience.

Experience-centred design (Wright & McCarthy, 2010) is an emergent field that seeks to combine all of the above elements with important key themes. The experience occurs firstly *in context*. The ways humans behave and perform actions are shaped by the environment and therefore this context must be a part of the design process. Experience-centred design is *participatory*, and this approach introduces a democratisation to the design process. Finally, experience-centred design is *humanist* and therefore emotional and should not be deconstructed.

“The point is that if you divide and separate emotion and behavior in this way you actually destroy the phenomenon you are seeking to study” (Wright & McCarthy, 2010, p. 14)

Experience-centred design has developed away from its original positivist scientific base of cognitive psychology to an interpretivist approach, focussed on the entirety of the encounter.

These approaches however, do not provide direct information concerning the variables that will make public service organisations effective. Instead they are focussed on methods and processes: the “operational management” (Osborne, et al., 2013) rather than the “design rules” that must be applied to enable the flexibility that the authors profess. Alford and Hughes (2008, p. 141) admit that developing such design rules are required but to date has not been addressed in the literature:

“There is therefore a need for a series of “design rules”—propositions about which managerial devices are called for in which situations. We do not propose to offer such a schema here...”

There is a definite reluctance in the literature to define such “design rules”. Some readdress the responsibility to the individual designers or engineers of services and suggest that the rules should be developed according to the services and products being developed (Polaine, Løvlie, & Reason, 2013, p. 152). Although this approach is useful in creating specific objectives and measures for each project, it does not provide a well thought out set of guidelines or typology that can be used to assist in comparing and measuring the effectiveness in designing public services. It also means that designers are recreating knowledge each time a new project takes place, rather than building on existing knowledge.

Others rely heavily on the use of SERVQUAL, the simplified RATER (Parasuraman, 1991; Parasuraman, et al., 1988) or the Triple Bottom Line (Elkington, 1997). As mentioned previously, SERVQUAL is inappropriate as a tool for measuring effectiveness in regards to behavioural intention. The Triple Bottom Line does have advantages in providing global outcomes measurement indicators with regards to effectiveness of public services. Rather than measuring profit margins and customer opinions it provides a framework for assessing the economic, environmental and social aspects of a service. However, again in the context of understanding and improving behavioural level AE it is not at the necessary level of specificity as the measures are at a macro or organisational level.

Information technology (IT) and service design

The information technology literature provides potential insights for measuring AE. The IT paradigm is moving away from what was essentially a “product-manufacturing dominant” approach to the “service dominant” approaches noted above. IT development relies heavily on Service Oriented Architecture (SOA) (Allen & Higgins, 2006, pp. 138-142). SOA is an evolution of enterprise architecture approaches. Although touting that SOA encompasses the full service experience, it is very much focussed on the tangible interfaces customers have with services, namely IT systems. In terms of quality measures, SOA is concerned with: *agility, capacity, availability; and security*. Design rules consist of elements related to the infrastructure such as ‘device independence’, ‘service dependency’, ‘layering’ and ‘data management’. Design guidelines include: coupling and cohesion, communication and agility (Allen & Higgins, 2006, p. 148).

Although the IT domain claims greater customer orientation, it is clear that it is still very much (software) product focussed. The language used throughout the literature causes confusion in this regard as a ‘service’ refers to a ‘software service’. Interface is less about the interaction between the customer and the software, and more about the interaction between two pieces of software.

A potential connection exists between AE and the Unified Theory of Acceptance and Use of Information Technology (Venkatesh, Morris, Davis, & Davis, 2003). The model replicates all of the variables in the Theory of Planned Behaviour with the exception of

“attitude towards the behaviour”. The additional variable ‘facilitating conditions’ focuses on the product use, not the overall service experience. Therefore, the application of this variable would be inadequate as a measure of AE.

Given the heavy reliance on information technology, the singularly inward-looking view of the IT literature domain raises concerns in terms of the broader services management and new public governance approach. The services and public governance literature is directed at understanding and designing holistic services and co-production, and the customer is central to the service delivery. Software development and general information technology development are no doubt important parts of this landscape. However, the IT literature is narrowly focussed on specifications, scope and detailed architecture, security, agility and data management. This reclusive vision does not lend itself to expansion and integration with the broader services picture. Human factors engineering is a more appropriate connection as it attempts to connect the engineering components of IT and the “softer” domain of psychology. This expansion includes the areas of Human-Computer Interaction and Human Factors engineering.

Ergonomics, Human Factors and usability

Human Factors and Ergonomics (HFE) emerged as a post-war necessity in terms of industrial design and engineering (Meister, 1999). During the 1940s and 1950s manufacturing and engineering focussed on automation, greater efficiency and swift instruction for tool and system handling as a result of the labour shortage. Aircraft production changed significantly to include rigorous testing to reduce accidents due to pilot error, which was a costly and in some cases, fatal consequence of poor design. The considerable interest of the defence forces on ‘error proofing’ their equipment meant that a number of international standards were developed (Military Standard MIL-STD 1472)¹¹. Extensive research followed, focussing on safety, fatigue and the effect of altitude, design of controls, cockpits and pilot performance (Nickerson, 1999).

Concurrently, the European Productivity Agency (EPA) launched a program of “making the job fit the worker” which ran contrary to the predominant thinking of cost efficiency (International Ergonomics Association, 2006). From the program emerged a field which was a surprising union between an interest in physical objects (engineering) and human performance of behaviour (psychology) (Meister, 2000). Ergonomics was propelled by necessary workforce improvements, job adaptability and a requirement to integrate those left with post-war disabilities back into the workforce.

¹¹ MIL-STD-1472F, DEPARTMENT OF DEFENSE DESIGN CRITERIA STANDARD: HUMAN ENGINEERING (23 AUG 1999). This standard establishes general human engineering criteria for design and development of military systems, equipment, and facilities. Its purpose is to present human engineering design criteria, principles, and practices to be applied in the design of systems, equipment, and facilities (Department of Defence, 1999)

Interest in the new field grew through the 1950s and 1960s with the foundation of the Human Factors and Ergonomics Society in 1957 (Human Factors and Ergonomics Society, 2015) and the International Ergonomics Association in 1959 (International Ergonomics Association, 2006). The field evolved through the influence of service delivery design and the need to create interfaces for increasingly complex information technology.

The majority of research in this field is focussed on insights and extensions from the cognitive sciences, and the development of research and design methods from engineering (Nickerson, 1999). However, a few general perspectives can be used to examine the quality of the outcomes based on design principles or heuristics in the field. McCormick provided an initial scheme of 18 design principles in 1964 and these have been republished in seven revisions of the standard text (McCormick, 1964; Sanders & McCormick, 1998). The seventh edition has the design elements organised according to the following categorisations: information input; human output and control; workplace design; environmental conditions and applications (Sanders & McCormick, 1993). The original publication was one of the first attempts to provide guidance to a field that was rapidly evolving. The principles can be used to consider control actions, information reception, decision making and general design for ergonomic psychology and engineering.

Universal Design (Erlandson, 2008; Gassmann & Reepmeyer, 2011) is a methodology that can be applied to all products and services. This approach was developed alongside mainstream human factors research; however it has a focus on accessibility and equity in design. This methodology is now considered in the development of a broad range of products (Gassmann & Reepmeyer, 2011). Principles of Universal Design include *equity, flexibility, simplicity, perceptibility, error recovery, low effort and accessibility* (Farage, Miller, Ajayi, & Hutchins, 2012).

Although the principles have been developed with the input of a range of design, ergonomics and engineering professionals, little empirical evidence demonstrates the internal validity of the principles and whether they are representative of the most critical universal design components (Story, 1998; Story, Mueller, & Montoya-Weiss, 2000; Story, Mueller, & Montoya-Weiss, 2001).

The narrow field of independent evidence for evaluation of universal design measures relates specifically to the effectiveness of the design measures in disability research. Beecher and Paquet (2005) developed a survey instrument to assist in applying universal design and usability principles broadly to consumer products. This instrument was used to evaluate with 36 disabled adults the use of four product families to determine their usability and accessibility. Statistical analysis showed that 11 factors were aligned with the seven principles of universal design: Low physical effort, perceptible information, flexibility in use, adaptability to user pace, intuitive use, reach and access of use, simple use, equitable use, secure, safe and private use, tolerance for error and size for use.

Lenker, Nasarwanji, Paquet and Feathers (2011b) used the research conducted by Beecher and Paquet (2005) to create a 12 item usability assessment tool known as the Rapid Assessment of Product Usability and Universal Design (RAPUUD). The authors made substantial inroads into practical application and evaluation of design with the development of this tool. However, they also admit further work is required to ensure the reliability and validity of the survey instrument. Other noted issues were flexibility and adaptability of the survey for either novice users or for those products that may require specific functionality due to additional user disabilities. Although not mentioned by the authors, a further limitation is the product-dominant focus of the research. Many of the products tested are simple use products: cabinets, showers/bathtubs, and stoves/ovens. Computers were also evaluated. Computers and stoves/ovens received low ratings for cognitive effort, which were the most complex products. However, these products were not tested in the context of performing a number of related tasks to complete an overall objective. Therefore, interactions with other products as well as temporal issues were not explored.

Usability, which is a predominant component of universal design and the user-centred design methodologies, has a set of design principles mainly associated with the design of websites and graphical user interfaces, but which are also applied to general development of people-product interfaces. The usability principles according to Nielsen (1996) are: *make screens simple and natural, speak the users language, minimize memory load, be consistent, provide feedback, provide clearly marked exits, provide shortcuts for advanced users, use plain language for error messages, prevent errors, make help simple and task-focused*. A fundamental issue to understanding and applying usability is the inability of users to self-assess 'usability'. A survey instrument that is completed by the users of the products without any triangulated information from expert or critical observations may provide a misleading and biased assessment of the product's usability (Nielsen, 1994, p. 12).

A number of professionals in the field of human factors and ergonomics have called for more work to be done around effectiveness assessment (Committee on Human Factors, 1985; Meister, 1986) particularly in regards to demonstrating the value of human factors.

This requires a more systematic approach to design quality and usage issues than is typically the case. In other words design quality and usage issues need to be incorporated into the specific design objectives that product developers have to meet (Wilson & Corlett, 2005, p. 322).

Although some researchers believe that Human Factors intrinsically assesses the "goodness" of the design as a natural part of the methodological approaches, these evaluations are usually specific to a product or system (Sanders & McCormick, 1998, p. 25). Quality control is predominantly related to the method applied rather than specific measures. Such methods include Activity Theory (Decortis, Noirfalise, & Saudelli, 2000; Gould & Verenikina, 2003; Seglin, Bedny, & Meister, 2000), Human-

computer Interaction (HCI)(Hornbæk, 2006) and User-Centred Design (UCD)(Durmus, 2012). The field is also governed by international standards¹² which provide quality control around the design and research techniques.

AE has many shared characteristics with the human factors and ergonomics domain. Principally, AE is directed at understanding and overcoming the obstacles to performance. Error prevention and recovery are critical considerations for both AE and HFE. One of the principal researchers of this domain Norman (1990, p. viii) describes the core premise of understanding human factors:

“Most accidents are attributed to human error, but in almost all cases human error was the direct result of poor design. The principles that guide a quality, human-centred design are not relevant just to a more pleasurable life – they can save lives.”

Norman (1990, p. viii) described a typical real life example of how error prevention is a primary design concern using the Three Mile Island disaster in the United States.

“I was among a group of social and behavioural scientists who were called in to determine why the control-room operators had made such terrible mistakes. To my surprise we concluded that they were not to blame: the fault lay in the design of the control room. Indeed the control panels of many power plants looked as if they were deliberately designed to cause errors.”

The work of HFE researchers has shown that many accidents ranging from simple and frustrating errors to fatalities are often the consequence of poorly designed or implemented systems. These outcomes are often the result of a lack of understanding of the real world use of the design subject and the potential for human error. Many limitations exist in human behaviour which may cause inadvertent error. These limitations relate to attention (selective, focussed, divided); sensation; vision (size, brightness, visual field, visual acuity); hearing; kinaesthetic; vestibular; physiology; smell and taste; strength; fatigue; perception; cognition and memory (Nemeth, 2004). The influence of external stressors will also impact negatively on the performance of behaviour. External stressors include noise, vibration, temperature, illumination, and sleep loss. Performance may also be affected by internal considerations or processes such as anxiety, depression, transient emotions and phobias (Nemeth, 2004).

Human Reliability Assessment theory (Spurgin, 2010), an emerging field of human factors, focusses on understanding how humans can be set-up to succeed or fail in task performance. Human failure probability assesses the human impact rather than the failure of a component of the system. 134 methods can be used to understand human failure in various medical systems (Lyons, 2005). However, no such equivalent methods exist in examining the role of the client-citizen experience for public administration and therefore consequently, no related evaluation methods for citizen-taxpayers in the use of tax administration systems. The literature clearly

¹² ISO13407:1999 Human-centred design processes for interactive systems and later replaced by ISO 9241-210:2010 Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems

recognises the complexity of tax administration but no known approaches to identifying and eliminating such errors for taxpayers exist. This is a clear gap in the design and evaluation of public services such as tax.

Error prevention mechanisms would provide significant improvement in the design and evaluation of public service delivery. Therefore error management is pertinent to the development of variables for AE. The tax system and the services with which taxpayers must interact are especially complex, ambiguous and resource intensive. Therefore, understanding and controlling for human errors in AE is imperative.

5. Constructing Administrative Effectiveness from design principles

The literature indicates that there are a number of dimensions potentially applicable to the construction of AE. Subsequent to the literature review, the primary variables used in each of the domains and theories were compared. A comprehensive cross matching of the dimensions and elements listed in the literature was first performed. The factors were then categorised and evaluated to reduce duplication. The reduced set of factors were then assessed for their contribution to a service-dominant understanding of public administration, role in co-production and similarities with user-centred design: all aspects considered necessary in Administrative Effectiveness as deduced from the above literature. The factors were then further categorised into relevant themes to develop the proposed model.

The principles for HFE and for Universal Design are the most comprehensive from the range of literature discussed and are also aligned with AE. Therefore, these principles were used as the base comparison set for the other domains. The principles were evaluated to assess their similarity or difference to a corresponding value in Universal Design. Appendix A lists the core theories and the related principles. Appendix B shows a summary of each of the key design and evaluation principles from the two most relevant domains: service management and HFE. The original factors from SERVQUAL were used to compare against those of Universal Design, rather than those from subsequent studies. Many of the items in the short versions of SERVQUAL in later studies were based on untested factors. The original complete set of variables also allows for greater granularity and comparison.

When compared against the usability heuristics the only variable that was not catered for already in universal design was 'Aesthetics'. Aesthetics relates to the visual appeal of the object or objects with which the individual interacts.

A number of missing constructs were evident in the cross comparison against SERVQUAL. Responsiveness, competence, credibility and courtesy are all characteristics of an interpersonal interaction (although some systems are anthropomorphised to demonstrate these characteristics). It is plausible that these components would not be evident in universal design variables because universal design developed out of a product-dominant focus and is predominantly interested in a task-specific interaction. Security was also absent from the universal design

variables. The inclusion of these variables will add to the evaluation of the humanistic side of the AE.

There were also a number of apparent gaps in the literature where no existing dimensions existed to measure a relevant construct. These areas include: usefulness of service, awareness, completeness or seamlessness of the experience, complexity or simplicity of the experience, the effect of context on the experience, temporal considerations, the interactive effect between the participants on future expectations and emotion.

The final set of potential dimensions to measure AE are shown in Table 1 with the following primary categories: people-product; people-process; and people-service.

Table 1 - Final set of variables to be used to measure AE

| People-Product | People-Process | People-Service |
|-----------------------|------------------------------|--------------------------|
| Cognition | Flexibility | Responsiveness (two way) |
| Perception | Error management | Competence |
| Ergonomics | Efficiency | Courtesy |
| Aesthetics | Stability and predictability | Credibility |
| Accessibility | Security | Equity |
| Usefulness | Complexity/simplicity | Co-production |
| Awareness | Completeness/seamlessness | Interactive expectations |
| | Context | Emotion |
| | Temporal considerations | |

Additionally, the literature did not provide connection to the outcome of an improved client experience or the role that a well-designed service has on willing participation or compliance. These elements must also be considered in a model for administrative effectiveness.

6. Discussion and conclusions

The literature review demonstrates the breadth of research concerned with the evaluation of design and service quality in a number of domains including services management, public administration and human factors. Public administration and its more recent conceptualisation, New Public Management has developed in seclusion from the service design, marketing and HFE literatures. The lack of augmented collaboration across these fields has meant that public management has evolved with product dominance and a focus on satisfaction and productivity measures. Although having advanced significantly with the introduction of concepts of coproduction, a complete understanding of the quality of the service delivery from the client perspective remains elusive.

All of the disciplines discussed have conceptually distinct origins but are converging around a new idea of client-citizen-consumer experience. By utilising research and knowledge from multiple sectors, a robust conceptualisation of the AE construct can be achieved. The diverse areas in the literature provide general principles that must be considered in the development of a new approach to evaluating the quality of a public service through the client-citizen experience: the AE variable.

Advances in evaluating and therefore improving public services must come from a synthesis of the knowledge gained from a number of disparate areas. The evolution of public administration or services management has created a divergence in understanding of service quality. Consequently, there has been a narrowing of focus, ongoing introspective re-examination of the methods and principles that are indigenous to each area of the literature. The public administration discipline must be invigorated through incorporation of ideas from respective fields, each of whom are contending with similar issues around designing and evaluating the client-citizen experience. Many domains have made important progress in specific focal areas. These advances must be shared and consolidated to provide a quantum leap in the complex area of administration and delivery of public services. It is necessary to take the opportunity to assess each and determine the value that they would bring to a new conceptualisation of service quality specifically for public administration.

This new model for public service design quality includes a number of core constructs. Both services and products must be included and designed into a client experience. Designing products in isolation from the full service means that the context of use and the possibility for errors created by external factors are not considered and therefore are not mitigated in the design. Any new model for evaluating and designing improved public services must include a method for identifying, and if possible, eliminating unintentional performance errors.

Additionally, the client-citizen-consumer experience is constructive and evolutionary. Interactions between the 'user' and the 'producer' are not discrete and are in fact co-produced in a form of exchange. Co-production is necessary to improve service quality and consumer satisfaction. Digitisation and information technology present additional challenges in ensuring a seamless and integrated experience. Information technology has developed around systems with a "product-focus" often at the expense of a broader customer experience view.

We highlight the need to rebalance the role of individual products and how they are combined in a view of the client-citizen experience. Important, but often over-valued products such as IT or digital applications must be considered in the context of the broader service ecology. An ongoing focus on IT solutions or products can often occur at the expense of the broader client experience. Figure 1 depicts a general model for future research to facilitate the exploration of AE. This model shows a rebalancing of the product with the service and process in terms of a complete picture of AE. The juncture of these elements determines the client-citizen experience. When

these elements occur in equilibrium, the experience will be optimised. Administrative Effectiveness will provide an overall measure of the quality of the experience.

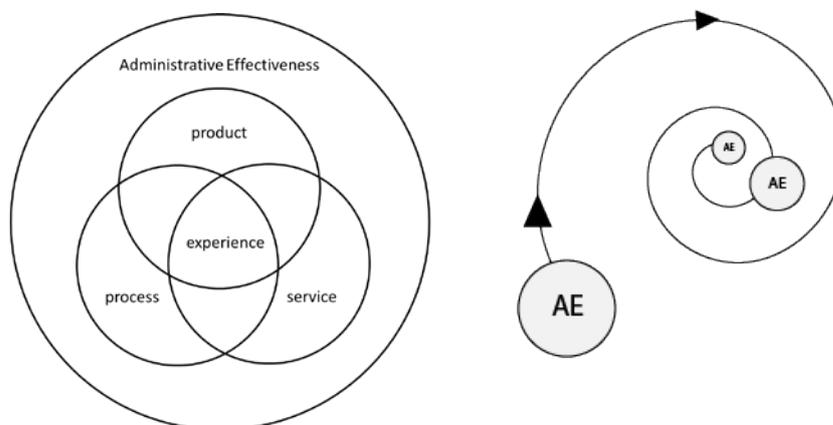


Figure 1: Model for Exploring AE

Figure 2: The iterative process for improving the client experience of the created through co-production of the service

Figure 2 depicts the aspect of the client-citizen experience determined by the effect of coproduction. Coproduction forces the experience to change for each participant in the experience. Each time individual actors participate in the experience the experience is refined or shifted, building on knowledge, satisfaction and associated emotions from the previous experience. Ideally, this experience is improved or refined with each iteration.

When Administrative Effectiveness is high, improved compliance should be a resultant outcome. High AE will also result in a positive, good or “high” client experience.

The model itself must also be considered as dynamic. As AE increases, an ongoing effect on the long term client experience should result such that high levels of AE will shape the future client experience in a positive direction. Client-citizens will be seen as participating in the service, providing input and feedback as well as setting ongoing expectations of the system. The system will respond through improvements in the service: directional changes to match the respective expectations of the client-citizens. Staff who are interacting with the client-citizens will also participate in this ongoing evolution through positive interactions and greater cooperation, service delivery and discretionary effort. It is hypothesised that modifications to the Administrative Effectiveness of a public service would not be linear but would be shaped by ongoing influences of the context, environment and the participants (both staff providing the services as well as the client-citizens). This evolution would continue to a point where a completely new service appears. The new service would then proceed through the same evolutionary process.

Concluding summary

In the private sector, the *client-customer* experience is predominantly defined and measured according to customer satisfaction and feedback. Additional success measurement is through profit and increase in market share for the business. Although this approach has limitations, it has been an effective method for approximating the effectiveness of a large range of service experiences. Human Factors and Ergonomics has been refining the design and development of human-user experiences, primarily focussed at workplaces with high risk activities (such as piloting aircraft) to reduce accidents and minimise human error. Success for HFE is founded in the prevention of fatalities. This domain provides significant insights into methods for improving client-citizen experience designs. However, neither of these domains have effectively incorporated the complexity of the *client-citizen* experience which is now anticipated in the New Public Administration. Service delivery and success depends on the cooperative relationship between the public administration and the citizens. Good public service design must deliver all of the aforementioned as well as create mutual trust and participation. Co-production and the constant evolution of the client-citizen-experience is an important facet of the development of public service design.

Each of the noted research domains have unique contributions, which when combined will provide insights and further opportunities for extension of the respective fields. Improving tax compliance behaviour is both a relevant and a multifaceted issue. The research in the field of tax compliance already extends across the domains of economics, sociology, criminology and psychology. However, improvements to compliance behaviour may only come from a nexus of an unlikely partnership with domains not normally associated with tax compliance: service quality and HFE. Future research is necessary to evaluate and estimate the model and the dimensions proposed in this paper.

7. References

- Ajzen, I. (1991). The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211. doi: 10.1016/0749-5978(91)90020-t
- Alexandria, B., & Adrienne, C. (2001). Service improvements in public services using SERVQUAL. *Managing Service Quality*, 11(6), 389.
- Alford, J. (2002). Defining the client in the public sector: A social-exchange perspective. *Public Administration Review*, 62(3), 337-346. doi: 10.1111/1540-6210.00183
- Alford, J., & Hughes, O. (2008). Public value pragmatism as the next phase of public management. *The American Review of Public Administration*, 38(2), 130-148. doi: 10.1177/0275074008314203
- Allen, P., & Higgins, S. (2006). *Service orientation: Winning strategies and best practices*. New York: Cambridge University Press.
- Ashworth, R., Boyne, G., & Entwistle, T. (2010). *Public service improvement: theories and evidence*. New York: Oxford University Press Oxford.
- Australian Taxation Office. (2009). *ATO Annual Report 2008/09*. Canberra: Retrieved from www.ato.gov.au/annualreport.
- Australian Taxation Office. (2013). Reinventing the ATO – ready, willing and underway, from <http://www.ato.gov.au/Media-centre/Speeches/Reinventing-the-ATO---ready,-willing-and-underway/>
- Australian Taxation Office. (2014a). Audit effectiveness: Impacts of audits on taxpayer behaviour and revenue. Canberra: Australian Taxation Office.
- Australian Taxation Office. (2014b). Preventing debt: Personas and pathways. Canberra: Australian Taxation Office.
- Australian Taxation Office. (2015). *Reinventing the ATO: program blueprint March 2015*. Canberra: Commonwealth of Australia.
- Barnard, C. I. (1938). *The functions of the executive*. Cambridge Harvard university press.
- Beccaria, C., & Voltaire. (1801). *Essay on crimes and punishments*. London: Printed by E. Hodson for J. Bone.
- Beck, P. A., Rainey, H. G., & Traut, C. (1990). Disadvantage, Disaffection, and Race as Divergent Bases for Citizen Fiscal Policy Preferences. *The Journal of Politics*, 52(1), 71-93. doi: 10.2307/2131420
- Beecher, V., & Paquet, V. (2005). Survey instrument for the universal design of consumer products. *Applied Ergonomics*, 36(3), 363-372. doi: <http://dx.doi.org/10.1016/j.apergo.2004.10.014>
- Bevacqua, J. (2015). ATO accountability and taxpayer fairness: An assessment of the proposal to split the Australian taxation office. *University of New South Wales Law Journal, The*, 38(3), 995-1014.
- Blanthorne, C., & Yance, K. S. (2006). The tax gap: Measuring the IRS's bottom line. *The CPA Journal* (April). Retrieved from <http://www.cpapjournal.com/2006/406/essentials/p40.htm>
- Blumer, H. (1969). *Symbolic interactionism: Perspective and method*. Englewood Cliffs: Prentice-Hall.
- Braithwaite, V. (2009). *Defiance in taxation and governance: Resisting and dismissing authority in a democracy*. Cheltenham, Glos: Edward Elgar Publishing Ltd.
- Chou, D. C., & Chou, A. Y. (2000). A guide to the Internet revolution in banking. *Information Systems Management*, 17(2), 51.
- Committee on Human Factors, N. R. C. (1985). *Methods for designing software to fit human needs and capabilities: Proceedings of the workshop on software human factors*: National Academies Press.
- Commonwealth of Australia. (2014). Budget paper no. 2: Budget measures Retrieved 18/2/2015, 2015, from http://budget.gov.au/2014-15/content/bp2/html/bp2_expense-22.htm
- Commonwealth Ombudsman. (2014). *Commonwealth Ombudsman Annual Report 2013-14*. Canberra: Commonwealth of Australia.
- Congram, C., & Epelman, M. (1995). How to describe your service. *International Journal of Service Industry Management*, 6(2), 6-23. doi: 10.1108/09564239510084914
- Dabholkar, P. (1995). The convergence of customer satisfaction and service quality evaluations with increasing customer patronage. *Journal of consumer satisfaction, dissatisfaction and complaining behavior*, 8, 32-43.
- Davies, M. (1998). Civil servants, managerialism and democracy. *International Review of Administrative Sciences*, 64(1), 119-129. doi: 10.1177/002085239806400109
- de Ruyter, K., Bloemer, J., & Peeters, P. (1997). Merging service quality and service satisfaction. An empirical test of an integrative model. *Journal of Economic Psychology*, 18(4), 387-406. doi: [http://dx.doi.org/10.1016/S0167-4870\(97\)00014-7](http://dx.doi.org/10.1016/S0167-4870(97)00014-7)
- Decortis, F., Noirfalise, S., & Saudelli, B. (2000). Activity theory, cognitive ergonomics and distributed cognition: Three views of a transport company. *International Journal of Human - Computer Studies*, 53(1), 5-33. doi: 10.1006/ijhc.2000.0378
- Deming, W. E. (1986). *Out of the crisis*. Cambridge, Mass Massachusetts Institute of Technology, Center for Advanced Engineering Study.
- Department of Defence. (1999). *Military Standard MIL-STD 1472 Human Engineering*: Department of Defence, United States of America.
- Durmus, S. (2012). *User-centered design approach in e-government applications*. Dissertation/Thesis, METU.

- Elkington, J. (1997). *Cannibals with forks: The triple bottom line of 21st Century business*. Oxford: Capstone.
- Erlanson, R. (2008). *Universal and accessible design for products, services, and processes*. Boca Raton, FL: CRC Press.
- Farage, M., Miller, K., Ajayi, F., & Hutchins, D. (2012). Design principles to accommodate older adults. *Global journal of health science*, 4(2), 2. doi: 10.5539/gjhs.v4n2p2
- Farnham, D., & Horton, S. (1993). *Managing the new public services*. Basingstoke, England: Macmillan.
- Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. New York: Psychology Press (Taylor & Francis).
- Frederickson, H. (1997). *The spirit of public administration*. United States.
- Funnell, S. (1997). Program logic: an adaptable tool for designing and evaluating programs. *Evaluation news and comment*, 6(1), 5-17.
- Gabbott, M., & Hogg, G. (1998). *Consumers and services*. Chichester; New York: J. Wiley.
- Gallagher, M. (2005). Benchmarking tax systems. *Public Administration and Development*, 25(2), 125-144. doi: 10.1002/pad.353
- Garvin, D. A. (1983). Quality on the line. *Harvard Business Review*, 61(5), 6-75.
- Gassmann, O., & Reepmeyer, G. (2011). Universal design: Innovations for all ages. In F. Kohlbacher & C. Herstatt (Eds.), *The silver market phenomenon* (pp. 101-116). Berlin, Heidelberg: Springer.
- Goldkuhl, G., & Lind, M. (2009). Grounding business interaction models: Socio-instrumental pragmatism as a theoretical foundation *Human Computer Interaction: Concepts, Methodologies, Tools, and Applications* (pp. 2071-2089): IGI Global.
- Golembiewski, R. T. (1969). Organization Development in Public Agencies: Perspectives on Theory and Practice. *Public administration review*, 29(4), 367-377.
- Gould, E., & Verenikina, I. (2003). *An activity theory framework for computer interface design*. Paper presented at the 25th International Conference on Information Technology Interfaces, 16-19 June.
- Gray, A., & Jenkins, W. (1986). Accountable management in British central government: Some reflections on the financial management initiative. *Financial Accountability and Management*, 2(3), 171-186.
- Gronroos, C. (1978). A service-oriented approach to marketing of services. *European Journal of Marketing*, 12(8), 588.
- Hasseldine, J., Evans, C., Hansford, A., Lignier, P., Smulders, S., & Vaillancourt, F. (2012). *A comparative analysis of tax compliance costs and the role of special concessions and regimes for small business in Australia, Canada, South Africa and the United Kingdom*. Paper presented at the National Tax Association Conference, Providence, Rhode Island.
- Heinrich, C. J. (2002). Outcomes-based performance management in the public sector: Implications for government accountability and effectiveness. *Public Administration Review*, 62(6), 712-725. doi: 10.1111/1540-6210.00253
- Heinrich, C. J. (2009). *Measuring Public Sector Performance and Effectiveness. Handbook of Public Administration*. SAGE Publications Ltd. London: SAGE Publications Ltd.
- Heskett, J. L., Jones, T. O., Loveman, G. W., Sasser, W. E., & Schlesinger, L. A. (1994). Putting the service-profit chain to work. *Harvard Business Review*, 72(2), 164-174.
- Hessing, D. J., Elffers, H., & Weigel, R. H. (1988). Exploring the limits of self-reports and reasoned action: An investigation of the psychology of tax evasion behavior. *Journal of Personality and Social Psychology*, 54(3), 405-413.
- Hoelzl, E., Kirchler, E., & Wahl, I. (2008). Enforced versus voluntary tax compliance: The "slippery slope" framework. *Journal of Economic Psychology*, 29(2), 210-225. doi: 10.1016/j.joep.2007.05.004
- Hood, C. (1991). A public management for all seasons? *Public Administration*, 69(1), 3-19. doi: 10.1111/j.1467-9299.1991.tb00779.x
- Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, 64(2), 79-102. doi: <http://dx.doi.org/10.1016/j.ijhcs.2005.06.002>
- Human Factors and Ergonomics Society. (2015). About HFES Retrieved 28/1/2015, 2015, from <https://www.hfes.org/web/AboutHFES/about.html>
- Hutchens, G. (2014, 15 July 2014). Fewer ATO jobs may cost \$1b in revenue, *Sydney Morning Herald*. Retrieved from <http://www.smh.com.au/federal-politics/political-news/fewer-ato-jobs-may-cost-1b-in-revenue-20140714-3bxab.html>
- Iacobucci, D., Ostrom, A., & Grayson, K. (1995). Distinguishing service quality and customer satisfaction: The voice of the consumer. *Journal of Consumer Psychology*, 4(3), 277-303. doi: 10.1207/s15327663jcp0403_04
- Inquiry into Tax Disputes*, (2014).
- Inspector-General of Taxation. (2005). Review into the Tax Office's Small Business Debt Collection Practices - summary of submissions.
- Inspector-General of Taxation. (2014a). *ATO administration of penalties*. Canberra: Commonwealth of Australia Retrieved from http://www.igt.gov.au/content/reports/ATO_administration_of_penalties/ATO_Administration_of_Penalties.pdf.
- Inspector-General of Taxation. (2014b). *Inspector-General of Taxation Annual Report 2013-14*. Canberra: Commonwealth of Australia.
- International Ergonomics Association. (2006). *IEA 50th Anniversary*: IEA Press.
- Jacobs, B. (2010). Deterrence and deterability. *Criminology*, 48(2), 417-441.

- Johnson, J. (2010). *Designing with the mind in mind: Simple guide to understanding user interface design rules*. Boston: Morgan Kaufmann.
- Jones, A. (2012). Where Has All the Public Administration Gone? *Teaching Public Administration*, 30(2), 124-132.
- Jordan, C. (2013). *Tax, the way ahead*. Paper presented at the Tax Institute Annual Convention.
- Jordan, C. (2014, 14/4/2014). *Reinventing the ATO: Building trust in Australia's tax administration*. Paper presented at the ATAX 11th International Tax Administration Conference, Sydney.
- Kaboolian, L. (1998). The New Public Management: Challenging the boundaries of the management vs. administration debate. *Public Administration Review*, 58(3), 189-193. doi: 10.2307/976558
- Kakouris, A., & Meliou, E. (2011). New Public Management: Promote the public sector modernization through service quality. Current experiences and future challenges. *Public Organization Review*, 11(4), 351-369. doi: <http://dx.doi.org/10.1007/s11115-010-0134-3>
- Kellogg, D., & Chase, R. (1995). Constructing an empirically derived measure for customer contact. *Management Science*, 41(11), 1734.
- Kirchler, E. (2007). *The economic psychology of tax behaviour*. New York: Cambridge University Press.
- Kirchler, E., Hoelzl, E., & Wahl, I. (2008). Enforced versus voluntary tax compliance: The "slippery slope" framework. *Journal of Economic Psychology*, 29(2), 210-225. doi: 10.1016/j.joep.2007.05.004
- Langham, J. (2012). *Good intentions aren't good enough: the role of administrative effectiveness in predicting tax compliance behaviour*. Master of Philosophy, University of Queensland, St. Lucia, Qld. (Book, Whole)
- Lenker, J., Nasarwanji, M., Paquet, V., & Feathers, D. (2011a). A tool for rapid assessment of product usability and universal design: Development and preliminary psychometric testing. *Work-a Journal of Prevention Assessment & Rehabilitation*, 39(2), 141-150. doi: 10.3233/wor-2011-1160
- Lenker, J., Nasarwanji, M., Paquet, V., & Feathers, D. (2011b). A tool for rapid assessment of product usability and universal design: Development and preliminary psychometric testing. *Work: A Journal of Prevention Assessment & Rehabilitation*, 39(2), 141-150. doi: 10.3233/wor-2011-1160
- Levine, C. H., Backoff, R. W., Cahoon, A. R., & Siffin, W. J. (1975). Organizational Design: A Post Minnowbrook Perspective for the "New" Public Administration. *Public Administration Review*, 35(4), 425-435.
- Loftus, E., & Palmer, J. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, 13(5), 585-589. doi: 10.1016/S0022-5371(74)80011-3
- Lyons, M. (2005). *Error reduction in medicine: Final report to the Nuffield trust*: Nuffield Trust.
- Maor, M. (1999). The paradox of Managerialism. *Public administration review*, 59(1), 5-18.
- Mauldin, C. (2013). Designing better public services. *Innovation Quarterly Journal*, 32(2), 32-35.
- McCarthy, J., & Wright, P. (2004). *Technology as experience*. Cambridge, MA: MIT Press.
- McCormick, E. (1964). *Human factors engineering*. New York McGraw-Hill.
- McKerchar, M. (2002). *The impact of complexity upon unintentional noncompliance for Australian personal income taxpayers*. Doctor of Philosophy, UNSW, Sydney.
- McKerchar, M. (2003a). *The impact of complexity upon tax compliance: a study of Australian personal taxpayers* (Vol. no. 39.): Australian Tax Research Foundation.
- McKerchar, M. (2003b). *The impact of complexity upon unintentional noncompliance for Australian personal income taxpayers*. Dissertation/Thesis. Retrieved from http://uq.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMwY2BQMLFMS7ZMNEgzSTJOMzM3TrEwSEwzNUozTbM0NTBJMkxBGetAKs3dhBiYUvNEGWTdXE0cPXRhRWN8Sk5OvDmwHJEa3dtKMBAAuwYp0owKCQbpJqZAOuz5CQTUxPjxBRgPw_YyUo1NUxKA7rAIhUA2fshqw
- Meister, D. (1986). *Human factors testing and evaluation* (Vol. 5.). Amsterdam: Elsevier.
- Meister, D. (1999). *The history of human factors and ergonomics*. Abingdon, Oxford: CRC Press Inc.
- Meister, D. (2000). Theoretical issues in general and developmental ergonomics. *Theoretical Issues in Ergonomics Science*, 1(1), 13-21. doi: 10.1080/146392200308444
- Midwinter, A. F. (1990). Editorial: What New Managerialism? *Public Policy and Administration*, 5(1), 2-3. doi: 10.1177/095207679000500101
- Miller, G. (1960). *Plans and the structure of behavior*. New York Holt, Rinehart and Winston.
- Millward Brown. (2014a). Attitudinal and behavioural research on the prevention of aged debt: Qualitative. Auckland: Millward Brown.
- Millward Brown. (2014b). *Perception of fairness in disputes: Wave 1 results*. Australian Taxation Office.
- Mo, T., Xu, X. F., & Wang, Z. J. (2008). A service behavior model for description of co-production feature of services. In K. Mertins, R. Ruggaber, K. Popplewell & X. Xu (Eds.), *Enterprise interoperability III: New challenges and industrial approaches* (pp. 247-260). London: Springer-Verlag.
- Mohamed, M. (2012). Estimating the underground economy from the tax gap: The case of Malaysia. *Malaysian journal of economic studies*, 49(2), 91-109.
- Murphy, K. R. (1995). *Understanding performance appraisal : social, organizational, and goal-based perspectives*. United States.
- Nemeth, C. P. (2004). *Human factors methods for design: Making systems human-centered*. Abingdon, Oxford: CRC Press.

- Nickerson, R. (1999). Engineering psychology and ergonomics. In P. A. Hancock (Ed.), *Human Performance and Ergonomics* (pp. 1-45). San Diego: Academic Press.
- Nielsen, J. (1994). *Usability engineering*. San Francisco, CA: Morgan Kaufmann Publishers.
- Nielsen, J. (1996). Usability heuristics. *Health Management Technology*, 17(11), 34-. doi: 10348025
- Nielsen, J. (2001). First rule of usability? Don't listen to users Retrieved 27/1/2015, from <http://www.nngroup.com/articles/first-rule-of-usability-dont-listen-to-users/>
- Niemirowski, P., & Wearing, A. (2003). Taxation agents and taxpayer compliance. *Journal of Australian Taxation*, 6(2), 166-200.
- Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, N.J: Prentice-Hall.
- Norman, D. (1990). *The design of everyday things*. New York: Doubleday.
- OECD. (2014). Increasing Taxpayers' Use of Self-service Channels. Paris: OECD.
- Osborne, S. (2006). The new public governance? *Public Management Review*, 8(3), 377-387. doi: 10.1080/14719030600853022
- Osborne, S. (2009). Delivering public services: Are we asking the right questions? *Public Money & Management*, 29(1), 5-7. doi: 10.1080/09540960802617269
- Osborne, S., Bovaird, T., Martin, S., Tricker, M., & Waterston, P. (1995). Performance management and accountability in complex public programmes. *Financial Accountability and Management*, 11(1), 19-37. doi: 10.1111/j.1468-0408.1995.tb00159.x
- Osborne, S., Radnor, Z., & Nasi, G. (2013). A new theory for public service management? Toward a (public) service-dominant approach. *The American Review of Public Administration*, 43(2), 135-158. doi: 10.1177/0275074012466935
- Parasuraman, A. (1991). Refinement and reassessment of the SERVQUAL scale. *Journal of Retailing*, 67(4), 420.
- Parasuraman, A., Zeithaml, V., & Berry, L. (1985). A conceptual-model of service quality and its implications for future-research. *Journal of Marketing*, 49(4), 41-50. doi: 10.2307/1251430
- Parasuraman, A., Zeithaml, V., & Berry, L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of quality. *Journal of Retailing*, 64(1), 12.
- Pinho, J., Macedo, I., & Monteiro, A. (2007). The impact of online SERVQUAL dimensions on certified accountant satisfaction: The case of taxation services. *EuroMed Journal of Business*, 2(2), 154. doi: 10.1108/14502190710826031
- Podsakoff, P., MacKenzie, S., Lee, J., & Podsakoff, N. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879-903. doi: 10.1037/0021-9101.88.5.879
- Polaine, A., Løvlie, L., & Reason, B. (2013). *Service design*: Rosenfeld Media.
- Pollitt, C. (1993). *Managerialism and the public services: cuts or cultural change in the 1990s?* Oxford ; Cambridge, Mass: Blackwell Business.
- . Public Administration. (2008). In W. A. Darity, Jr. (Ed.), *International Encyclopedia of the Social Sciences* (2nd ed. ed., Vol. 6, pp. 603). Detroit: Macmillan Reference USA.
- Qiu, R. (2014). *Service science: The foundations of service engineering and management*. Hoboken, NJ: John Wiley & Sons.
- Quelch, J., & Ash, S. (1981). Consumer satisfaction with professional services. In J. H. Donnelly, W. R. George & A. American Marketing (Eds.), *Marketing of services* (pp. 82-85). Chicago, Ill: American Marketing Association.
- Rothengatter, M. (2005). 'Sticks, carrots or sermons?' *Improving voluntary tax compliance among migrant small-business entrepreneurs of a multi-cultural nation*. Centre for Tax System Integrity, Australian National University. Canberra.
- Sampson, S., & Froehle, C. (2006). Foundations and implications of a proposed unified services theory. *Production and Operations Management*, 15(2), 329-343. doi: 10.1111/j.1937-5956.2006.tb00248.x
- Sanders, M. S., & McCormick, E. J. (1993). *Human factors in engineering and design*. New York, N.Y: McGraw Hill Book Company.
- Sanders, M. S., & McCormick, E. J. (1998). Human factors in engineering and design. *Industrial Robot: An International Journal*, 25(2), 153-153. doi: 10.1108/ir.1998.25.2.153.2
- Sanford, C. (1994). International comparisons of administrative and compliance costs of taxation. *Australian Tax Forum*, 11(3), 291.
- Schmenner, R. (1986). How can service businesses survive and prosper? *Sloan management review*, 27(3), 21.
- Schneider, F., Buehn, A., & Montenegro, C. (2010). *Shadow economies all over the world: New estimates for 162 countries from 1999 to 2007*. Working Papers wp322. University of Chile, Department of Economics.
- Seglin, M., Bedny, G., & Meister, D. (2000). Activity theory: History, research and application. *Theoretical Issues in Ergonomics Science*, 1(2), 168-206. doi: 10.1080/14639220050171324
- Shostack, G. (1982). How to design a service. *European Journal of Marketing*, 16(1), 49-63. doi: 10.1108/EUM000000004799
- Shostack, G. (1984). Designing services that deliver. *Harvard Business Review*, 62(1), 133-139.
- Spurgin, A. (2010). *Human reliability assessment: theory and practice*. Boca Raton: CRC Press.

- Story, M. (1998). Maximizing usability: The principles of universal design. *Assistive Technology*, 10(1), 4-12. doi: 10.1080/10400435.1998.10131955
- Story, M., Mueller, J., & Montoya-Weiss, M. (2000). *Progress in the development of universal design performance measures*. Paper presented at the RESNA 2000 Annual Conference.
- Story, M., Mueller, J., & Montoya-Weiss, M. (2001). *Completion of universal design performance measures*. Paper presented at the Proceedings of the RESNA 2001 Annual Conference.
- Stryker, S. (1980). *Symbolic interactionism: A social structural version*. Menlo Park, CA: Benjamin/Cummings.
- Sunstein, C., & Thaler, R. (2003). Libertarian Paternalism is not an oxymoron. *The University of Chicago Law Review*, 70(4), 1159-1202.
- The Australian Government the Treasury. (2015). *Re:think: tax discussion paper*. Canberra: Commonwealth of Australia.
- Unkelbach, C. (2006). The learned interpretation of cognitive fluency. *Psychological Science*, 17(4), 339-345.
- Vargo, S., & Lusch, R. (2004). Evolving to a new dominant logic for marketing. *Journal of Marketing*, 68(1), 1-17. doi: 10.1509/jmkg.68.1.1.24036
- Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User acceptance of information technology: Toward a unified view. [Article]. *Mis Quarterly*, 27(3), 425-478.
- von Hippel, E. (1994). "Sticky information" and the locus of problem solving: Implications for innovation. *Management Science*, 40(4), 429-439. doi: 10.1287/mnsc.40.4.429
- Wemmerlöv, U. (1990). A taxonomy for service processes and its implications for system design. *International Journal of Service Industry Management*, 1(3), 20-40. doi: 10.1108/09564239010002126
- Wilson, J., & Corlett, N. (2005). *Evaluation of human work*: CRC Press.
- Wright, P., & McCarthy, J. (2010). Experience-centered design: Designers, users, and communities in dialogue. *Synthesis Lectures on Human-Centered Informatics*, 3(1), 1-123. doi: 10.2200/S00229ED1V01Y201003HCI009
- Yang, Z., & Jun, M. (2008). Consumer perception of e-service quality: From Internet purchaser and non-purchaser perspectives. *Journal of business strategies*, 25(2), 59-84.

8. Appendix A

Effectiveness and design evaluation principles according to theoretical approach

| Experience centred design ¹³ | Universal design ¹⁴ | Usability Heuristics ¹⁵ | Usability Heuristics ¹⁶ | RAPUUD ¹⁷ | SERVQUAL ¹⁸ (reduced) | SERVQUAL (full) |
|---|-----------------------------------|---|---|----------------------|----------------------------------|------------------------------------|
| Compositional | Cognitively sound | Simple and intuitive use Make screens simple and natural | Visibility of system status | Ease of use | Reliability | Reliability |
| Sensual | Perceptible | Speak the users' language | Match between system and the real world | Mental effort | Responsiveness | Responsiveness |
| Emotional | Ergonomically sound | Minimise user memory load | User control and freedom | | Empathy | Competence |
| Spatio-temporal | Accessible | Be consistent | Consistency and standards | Physical effort | | Access |
| | Flexible | Provide feedback | Low physical effort | Information | | Courtesy |
| | Error managed/tolerance for error | Provide clearly marked exits | Error prevention/Tolerance for error | Mistakes and errors | | Communication |
| | Stable and predictable | Provide shortcuts for advanced users | | Time | | Credibility |
| | Equitable use | Use plain English for error messages | Recognition rather than recall | Safety | Security | Security |
| | Size/space for approach/use | Prevent errors | Flexibility and efficiency of use | Unwanted attention | | Understanding/knowing the customer |
| | | Make help simple and task focussed | Aesthetic and minimalist design | Embarrassment | Tangibles | Tangibles |
| | | | Help users recognise, diagnose and recover from error | Assistance | | |
| | | | Help and documentation | Clean up and storage | (plus additional) Convenience | |

¹³ (McCarthy & Wright, 2004)

¹⁴ (Erlandson, 2008)

¹⁵ (Gassmann & Reepmeyer, 2011)

¹⁶ (Nielsen, 1996)

¹⁷ (Lenker, Nasarwanji, Paquet, & Feathers, 2011a)

¹⁸ (Parasuraman, et al., 1988)

9. Appendix B

Matrix cross comparison of design theories to the Universal Design principles

| | Universal Design | People – user capabilities | | | Process | transcendental - holistic | | | | Other | |
|---|------------------|----------------------------|------------|------------|---------|---------------------------|------------------|------------|------------------------------|-------|--|
| | | Cognitive | perception | Ergonomics | | Flexibility | Error management | Efficiency | Stability and predictability | | Equitable |
| Usability | | | | | | | | | | | |
| Visibility of system status | | x | x | | | | | | | x | |
| Match between system and the real world | | x | x | x | | | | | | x | |
| User control and freedom | | | | | | x | x | x | | x | x |
| Consistency and standards | | | x | | | | | | | | x |
| Error prevention | | x | | | | | x | | | | |
| Recognition rather than recall | | x | x | | | | x | | | | x |
| Flexibility and efficiency of use | | | | | | x | | x | | | |
| Aesthetic and minimalist design | | | | | | | | | | | Missing aesthetics |
| Help users recognise, diagnose and recover from error | | | | x | | x | x | x | x | | |
| Help and documentation | | | | | | | x | | | | x |
| SERVQUAL | | | | | | | | | | | |
| Reliability | | | x | | | | | x | x | x | |
| Responsiveness | | | | | | | | | | | Missing responsiveness |
| Competence | | x | | | | | x | | x | | Missing competence |
| Access | | | | | | | | | | x | x |
| Courtesy | | | | | | | | | | | Missing courtesy |
| Communication | | x | | | | | | | | | |
| Credibility | | | | | | | | | | | Missing credibility |
| Security | | | | | | | | | | | Missing security |
| Understanding/knowing the customer | | | | | | x | x | | | | x |
| Tangibles | | | | | | | | | | | Missing tangibles –group with aesthetics |

10. Appendix C

Administrative Effectiveness dimension People-Product

People-product

Dimensions relate to the development of individual products within and as part of the interactions for the client-citizen experience

Cognition

Cognitive demands include, but are not limited to, memory, language, and learning requirements, as well as task complexity. (UD)

- Cognitively sound (UD)
- Simple and intuitive to use (UH)
- Visibility of system status (UH)
- Match between the system and the real world (UH)
- Provide feedback (UH)
- Speak the users' language (UH)
- Minimise user memory load (UH)
- Consistency and standards (UH)
- Recognition rather than recall (UH)
- Provide shortcuts for advanced users (UH)
- Ease of use (RAPUUD)
- Mental effort (RAPUUD)

Perception

Designed entities must effectively communicate necessary information to the user, regardless of ambient conditions or the user's sensory abilities (UD)

- Perceptible (UD)
- Be consistent (UH)
- Information (RAPUUD)
- Visibility (UH)

Ergonomics

Ergonomics refers to the physical demands of an activity, task, or job. (UD)

- Ergonomically sound (UD)
- Spatio-temporal (ECD)
- Low physical effort (UH)
- Physical effort (RAPUUD)

Accessibility

- Accessible design (UD)
- Access

Usefulness

- Needed
- Wanted
- Desired

Awareness

- Understood
- Comprehensible

11. Appendix D

Administrative Effectiveness dimension People-Process

People-process

Dimensions relate to the temporal considerations of the client experience, the order or sequence of activities and how the multiple parts of the interaction impact on each other

Flexibility

- Flexible (UD)
- User control and freedom (UH)
- Flexibility and efficiency of use (UH)
- Responsiveness (SERQUAL)

Temporal considerations

- Time
- Responsiveness (SERQUAL)
- Delays
- Sequence

Context

- The environment in which the experience takes place

Complexity/ simplicity

- The perceived difficulty of the process and interactions

Error management

- Error managed (UD)
- Tolerance for error (UD)
- Use plain English for messages (UH)
- Prevent errors (UH)
- Make help simple and task focussed (UH)
- Help users recognise, diagnose and recover from error (UH)
- Help and documentation (UH)
- Mistakes and errors (RAPUUD)

Efficiency

- Efficiency of use (UH)

Stability and predictability

- Stable and predictable(UD)
- Reliable (SERQUAL)

Security

- Safety (RAPUUD)
- Security (SERQUAL)

Completeness/ seamlessness

- Lack of fragmentation in experience

12. Appendix E

Administrative Effectiveness dimension People-Service

People-Service

Dimensions relate to the interaction of people participating in the overall experience.

Responsiveness

- Responsiveness (SERQUAL)

Competence

- Competence (SERQUAL)

Courtesy

- Courtesy (SERQUAL)
- Respect

Credibility

- Credibility (SERQUAL)

Equity

- Trust
- Fairness
- Consistency and standards (UH)

Emotion

- Sensual (ECD)
- Emotional (ECD)
- Empathy (SERQUAL)

Co-production

- Participation
- Past experience
- Future expectations

Interactive expectations

- Expectations