How Public Values affect civil servants’ perceptions on the success of Public Sector Reforms

Master Thesis

Martin Wasser

S1875000

BMS Faculty

Study program: PA Master

Supervisor: Dr. Veronica Junjan

Dr. Caroline Fischer

Profile Public Administration:

Regulation and Innovation

Word count: 13.744

20.03.2025

UNIVERSITY OF TWENTE.

Contents

Abstract

Aim: This research explores how public values affect German civil servants' perception of the success of public sector reform. Regarding the definition of perceived success, the academic community developed just a few approaches to what perception of success might entail, most of it in business management. Specifically, exploring factors that might determine how an individual views public sector reform. This thesis tries to approach this research gap with a special focus on the effect of public values on the perception of success.

Methods: Survey data collected in the EU COCOPS Program was used. This survey gathered perceptions about public sector reform by civil services and was conducted between 2012 and 2015. In the end, 4814 civil servants from 10 different European countries responded to the questionnaire that was sent. The sample used for this thesis included the answers of 445 respondents from Germany. This sample was analyzed using multiple regression analysis.

Results: The findings show significant, positive relationships between the public values of "productivity"; "and "performance" and the perception of the success of the public sector reform in the respondents' field. However, it was not possible to determine potential interconnections between public values and potential hierarchies of public values.

Implications: The findings do concur with past remarks by civil servants in terms of changes in the public sector. This shows that more coherent approaches to analyzing networks of public values and their interactions are needed. For the German public administrative system, the results imply that "classical" values that descend from Weberian principles are no longer as important as values that manifested under the NPM paradigm.

1. Introduction

The issue of the public sector and management reform is a topic that has kept scholars in the field of public administration occupied until today. There is an ongoing discussion about the public sector and civil service reform in OECD countries regarding digitalization such as Germany (Kaufmann, 2024) or Austria (Pitlik, 2017). Other such as the UK, engage more broadly in the general public sector reform (Feroz, 2023; Pitlik, 2017). Essentially the two latter cases have one thing in common: the resistance to reform comes from within the administrative body. In the UK, it has even gone so far as proposals for reforms, which have been sent to government public advisors have been disregarded or lower staff has been ordered to "bin it immediately" (Feroz, 2023). In the last example, public sector reform is discussed and planned without civil servants as stakeholders, but as enforcers or mere implementers. The academic literature on public sector reform reflects this disregard for civil servants as stakeholders, such as Pollitt & Bouckaert's state-of-the-art work "Public Sector Reform", which treats civil servants singularly as implementors rather than actual stakeholders. There seems to be a lower emphasis on how their "stake" might affect the outcomes of public sector reform. In their book, public sector reform is portrayed mainly on a macro-level: They put mechanisms of public sector institutions, internal reform processes, and national cultures as the main drivers for public sector reform, rather than focusing on the individual within the framework (Pollitt & Bouckaert, 2017, p. 33). Hence, it seems necessary to develop further approaches to better understand the actual stake that civil servants have in the public sector reform process.

The literature on public sector reform in recent years could be separated into different groups. With an overreaching focus on several aspects of governance for a specific country for example Japan (Ishihara, 2022) or Jordan (Alghizzawi, Masruki, & Auditing, 2020). Others try to deduct findings from the macro- to the meso-level of analysis by engaging in general issues of federal governance and the transfer into the meso-level by picking out specific agencies as the scope for the study (Alves, 2021). Interestingly though, there is a rise in literature on the mico-level and engagement with the roles of leadership, whether it concerns singled-out people (Bekelcha, 2023) or a specific group of civil servants (Ojogiwa, 2021).

Specifically, the last mentioned mico-level approaches are part of an ongoing development in recent years to analyze micro-level issues in public administration by including evidence from behavioral science and bringing them together with theories from public administration. This systematic approach of combining these two academic streams is collectively categorized under BPA (Behavioral Public Administration) (Bhanot & Linos, 2020). BPA scholars use theories and methods from psychology to analyze micro-level phenomena in public sector institutions. However, this field in public administration is not new as micro-level approaches can be traced back to Herbert Simon or Robert Dahl (Bertelli & M. Riccucci, 2022).

Since the beginning of the century, the discussion of values and their presence within the public domain and the constant change have been approached as an issue for the public administration research agenda, most famously by Jorgensen and Bozeman with their inventory of public values (2007). The concept of public values defines normative agreements between the citizenry and the government and the principles on which the government should design its policies (Bozeman, 2007). In more recent papers, the issue has been approached from different angles, as in the context of new technologies (Rodriguez Müller, Martin Bosch, & Tangi, 2024) or generally how to better asses what current public values are and how their respective importance is to be weighed against each other (Huijbregts, George, & Bekkers, 2022). In line with this revived discussion on what values are implicitly part of the public administration landscape and behavioral aspects of civil servants, the research question of this thesis is the following:

To what extent do public values explain differences in the perception of German civil servants regarding the success of reform in the administrative system?

In conclusion, this thesis will investigate if public values can affect a civil servant's belief that public sector reform in their field is successful. Concerning the structure of the thesis, the next chapter will deal with the main concepts and theoretical frameworks. Starting with a small literature review on reform change and reform evaluation, which will then move on to literature on the perception of reform success. Next, the thesis will move to the issue of how perception is formed on an individual level and what factors determine the formation of perception. The main concepts will be operationalized after developing hypotheses and presenting the theoretical model. Using the COCOPS dataset, the developed hypothesis will be tested. The thesis will end with a final chapter that includes the limitations of the thesis and further recommendations for practitioners and future academic discourses.

2. Theory

2.1 reform and reform evaluation

The literature on reform change in the public sector is a topic that has not lost its relevancy in contemporary academic literature. The only change concerns the scope of recent studies, which are predominantly focused on accounting and auditing reforms such as in Hungary (Lentner, Molnár, & Nagy, 2020), China (Zhang, 2021) or Uzbekistan (Abdullayeva, 2020). Some others are more concerned with the general change in developing countries (M'Mugambi, Okeyo, & Muthoka, 2020; Ukwandu, Ijere, & Review, 2020) While another group engages with reform change in the context of new technological developments (Collington, 2022; Park, Lee, Son, & Governance, 2021; Senyo, Effah, Osabutey, & Change, 2021). Moreover, with the rise of literature on reform change (as mentioned above) came a wave of attempts to revive the topic of policy evaluation in the public sector, something that has been coined by Pollitt (2003) as a "second coming" of evaluation studies (Pollitt, 2003). However, as Ugyel & O'Flynn point out present scholars of evaluation studies faced similar problems as their predecessors concerning a common rationale for evaluation to overcome systemic issues such as conflicting objectives of reform, lack of data, or inclusion of contextual factors (Ugyel & O'Flynn, 2017, p. 115). The authors further stress that reform evaluation is more disconnected from the evaluation process than policy evaluation, as practitioners limit the scope and sporadically execute the assessment. The lack of an independent institution that establishes evaluation criteria leads to a problem: There is no consistent definition of "goal achievement" that is beyond the influence of values and constructed facts (Ugyel & O'Flynn, 2017, p. 116). "Goal achievement" and thereby the perceived success is established by the authors themselves rather than an independent body and thus biased. Bovens and t'Hart therefore present two different kinds of groups of thoughts on how facts and values interact in policy evaluation:: The first group adheres to the rationalistic approach to evaluate policy by detaching values from facts to produce knowledge that is characterized as "apolitical" (Bovens, Hart, & Kuipers, 2006, p. 325). The second group advocates for a so-called "argumentative" policy evaluation. Meaning the assumption that facts cannot be detached from values. The evaluative process is defined by presenting the different views or realities that concerned actors have. ' There were attempts to bridge these two schools of thought by developing a "revisionist" approach to policy evaluation. First Boven and t'Hart followed a path where they did not create clear criteria of what is considered a policy success but measures for which a policy can be evaluated by determining the level of "programmatic" and "political" success. While programmatic success is determined by three criteria, namely effectiveness; efficiency, and resilience of the policy, "political success" is determined by the degree to which policies and policymakers are framed for their policies in the political arena. (Bovens, 2018, p. 656) Based on these two facets of policy success, Marsh and McConnell later developed their criteria of policy success by using Boven and t'Harts' foundation and developed three dimensions of policy success. The first two of "political" and "programmatic" success are equal in terminology to the ones by Bovens and that, however, the last one concerning "process" success deviates.

While the first three models incorporate concepts that are attributed to organizational concepts such as goals ("the goal model"), resources ("the systems-resource model") as well as processes ("Internal process model"), the latter two ("the competing values model" and "the multiple constituency model") deals with the predisposition of stakeholders and their interaction with policy evaluation. The issue, especially with the latter two, is that it still does not solve the problem of developing common indicators of success. Moreover, the multiple constituency model even rejects the possibility of common success indicators as "success" is defined by a group's chosen indicators to which they attach a weight (Marsh & McConnell, 2010, p. 567). This means that the effectiveness or success of an organization's output can be measured by the extent to which multiple constituencies perceive it as effective or successful (G. A. Boyne, 2003, p. 221). This model, therefore, displays that policy success is not defined by a settled set of criteria to which the responsible organization is seen as accountable, but rather by the struggle between different groups who evaluate outputs based on their inherited positions and values that they see as preferred. Thus, each actor gives a policy a different meaning based on their reality and this reality can be influenced by culture, power relations, values, etc. (Bovens et al., 2006, pp. 325-326). This is where authors that developed models and frameworks for policy evaluation in this context agree, namely that the perception of policy success is always connected to the struggle for power relations between the stakeholder groups (Bovens et al., 2006, p. 323; G. A. J. P. a. Boyne, 2003, p. 221; Marsh & McConnell, 2010, p. 576).

However, Marsh and McConnel have shortcomings in their theory by acknowledging what they frame as "complexity issues". These issues deal for once with the issue where the success of a policy is interpreted and perceived differently by the involved and affected stakeholders. The second issue deals with what the authors call "Time, space and culture", meaning the timespan of evaluation (which is shorter in politics than in the academic sphere), and on the other hand "space and culture" meaning variance of what is perceived as a success by different political entities (regions, nations, confederations, etc.) (Marsh & McConnell, 2010, p. 577). The same applies to George Boyne's work, as he acknowledges the strength of his models to display differences in interpretations of what states an effective (and thereby according to his logic a successful) organizational performance (G. A. Boyne, 2003, p. 220), but at the same time stresses that there is still a lack of understanding of how different public standards of a successful policy are related to each other and why they vary throughout involved stakeholders (G. A. Boyne, 2003, p. 223). Moreover, the stakeholders that assert more power over the criteria also determine the degree of success (Marsh & McConnell, 2010, p. 567).

To map out the different potential criteria that stakeholders might define, Byone (2003) developed the "competing values" model, where power over organizational activities plays a part in determining the success of reform. In this model, Boyne tried to lay out the contradicting criteria of performance public organizations face by categorizing these criteria into two different dimensions. One of these dimensions concerns the exchange between flexibility, where civil servants might have more discretion in the process, and the other with control where civil servants are closely monitored and scrutinized based on performance measures (G. A. J. P. a. Boyne, 2003, p. 219). The assessment of “control” can therefore be seen as a legitimate criteria that a stakeholder might use to evaluate the success of a reform.

In line with the argumentative school of policy evaluation, does the design of this model suggest that values in the context of public reform are best perceived by plotting one value against another in a dichotomous manner? It is Boyne’s way of portraying how different values interact with each other in the sphere of public policy and reform evaluation. This exchange helps overcome the previously mentioned issue concerning common evaluation criteria. Rather than developing criteria and arguing for them, the different values portray respective positions from which policies are evaluated from (G. A. J. P. a. Boyne, 2003, p. 224).

2.2 Public values and norms

The concept of public values is a topic in PA studies that is characterized as “unmanageable” due to its fundamental nature (Bozeman, 2007, p. 355). The concept was created as “ […] a rationale for government and other entities to defend and produce such values.” (Rainey, 2009, p. 64) Specifically, Bozeman presented in his book “*Public Values and Public Interest: Counterbalancing Economic Individualism”* (2007) a definition of what public values define: According to him, values are in a narrow sense public if “[…] those [provide] normative consensus about a) the rights, benefits and prerogatives to which citizens should (and should not) be entitled; (b) the obligations of citizens to society, the state, and one another; (c) the principles on which governments and policies should be based”(Bozeman, 2007).

One example of approaching the topic is Jørgensen & Bozeman’s typological of public values, which is independent of the entity that produces them (Rainey, 2009, pp. 67-68). In the approach by Jorgensen and Bozeman (2007), both authors presented a structured inventory of public values and categorized them into seven categories. Their inventory was based on a literature study of P.A. publications from 1990-2003 and specifically focused on literature from the U.S., UK, and Scandinavian countries (Jørgensen & Bozeman, 2007, p. 357). The result of this conducted research was an inventory that consisted of 72 registered values which they classified based on which public administration or organization is affected by the value. The authors ended up with seven different kinds of value categories:

Public sector's contribution to society

This category includes values associated with the idea that the public sector should contribute to society based on public interest. With a special focus on the concept of “public interest”, the public sector's contribution shall not happen according to special, individual interest but to support society as a whole rather than a particular group or class (Bozeman, 2007, p. 361). Other value groups of this category are “altruism” (acting in the interest of others); “sustainability” (awareness of the scarcity of resources and managing those resources for the sake of future generations) and “regime dignity” (the representation of a regime towards the outside world).

Transformation of interests into decisions

Moreover, the category about the transition from interests into decisions harbors values that are connected to principles such as the “majority rule” (democratic principles, the will of the people, and collective choice), the inclusion of the people into the democratic governance system (Local governance and citizen involvement) and lastly the protection of minorities (protection of individual rights) (Bozeman, 2007, p. 163).

Relationship between civil servants and politicians

This particular category entails the values of “accountability” and “responsiveness”. These values direct the relationship between public administrators and politicians in such a way that the emphasis is put on the politician, while the administrator acts as a utilizer and servant to the actions of the politician (Bozeman, 2007, p. 364).

Relationship between civil servants and their environment

The public values in this category are divided into three different subgroups, all designed dichotomously. The first one portrays the “openness” of a whole administration, whether it concerns the transparency of the internal decision-making process as well as the openness for input by the citizenry. On the opposite is the public value of secrecy, which means in the most autocratic sense the withholdment of information of interest to the public and in a democratic sense the protection of the personal data of the citizens(Jørgensen & Bozeman, 2007, p. 364). The second group entails values that define the priority between advocacy (for interests) and neutrality an administration tries to uphold in their decision-making. The last group is on the one hand the value of competitiveness, a value that originates from the private sector. Values that belong to this group are for example risk readiness or effectiveness. On the opposite of competitiveness stands cooperativeness, a value that is critical in the absence of market mechanisms where public organizations have to effectively cooperate to deliver services(Jørgensen & Bozeman, 2007, p. 366).

Intra-organizational aspects of public administration

This category contains four subgroups, namely “Robustness”; “Innovation”; Productivity” and “Self-development of employees”. The first subgroup is described by Jørgensen & Bozeman as a “machine metaphor” (Jørgensen & Bozeman, 2007, p. 366). These values deal with the sustainability of a public organization to function reliably, regardless of internal and external changes. The second group is about “Innovation”, which includes the values of “enthusiasm” and “riks-readiness”. These values define the degree to which public organizations can properly adapt to a dynamic environment by supporting a forward-thinking mindset, while “innovation” stays at the center of the three values. Connected to this is the third subgroup distinctively a collection of values that can be asserted with values of the New Public Management movement as “productivity”, “effectiveness”, “parsimony” and “business-like approach” are all associated with economic thinking(Jørgensen & Bozeman, 2007, p. 367).

Behavior of public-sector employees

In this category, all included public values are associated with the ideal envisions of civil servants, namely “professionalism”; “honesty”; “moral standards”; “ethical consciousness” and “integrity”. However, Jørgensen & Bozeman stress that the value of “Integrity” is considered to be the central value of this group, as a civil servant with high integrity is considered to act independently and thereby principled (Jørgensen & Bozeman, 2007, p. 368).

Relationship between public administration and the citizens

The last category includes four different subgroups. The first contains values associated with the legal status of a citizen and the relation to the state and includes values such as “legality”; protection of, the rights of the individual”; “equal treatment”, and “the rule of law”. This group is followed by the value subgroup of “equity”, which includes “reasonableness”; “fairness” and “professionalism”. This group is related to the first group because serving justice has to consider the citizen's circumstances. Additionally, the values of the equity subgroup have acted as guidance when civil servants make use of their discretionary power (Jørgensen & Bozeman, 2007, p. 369). The third subgroup deals with the interaction between civil servants and the citizens by considering values such as "responsiveness"; "user democracy"; "citizen involvement" and "citizen's self-development". All of these four values have their origin in the tradition of participatory democracy, where the civil servants receive their input from the citizens rather than deciding based on their own merits(Jørgensen & Bozeman, 2007, p. 369). The last subgroup is marked as "user orientation", which is a value group that is closely connected to private sector values of "friendliness" (The customer is king) as well as "timeliness" where the civil servants orientate towards the "customers" (citizens) need (Jørgensen & Bozeman, 2007, p. 369). Table (1) summarizes the public values that were sorted into one of the seven categories in the works by Jørgensen & Bozeman (2007).

Table 1

Table of value categories and their respective values according to Jørgensen & Bozeman

Value category	Included values
Public sector's contribution to society	Common good; Public interest; Social cohesion; Altruism; Human dignity; Sustainability; Voice of the future; Regime dignity; Regime stability
Transformation of interests into decisions	Majority rule; Democracy; Will of the people; Collective choice; User democracy; Local governance; Citizen involvement; Protection of minorities; Protection of individual rights
Relationship between civil servants and politicians	Political loyalty; Accountability; Responsiveness
Relationship between civil servants and their environment	Openness–secrecy; Responsiveness; Listening to public opinion; Advocacy–neutrality; Compromise; Balancing of interests; Competitiveness–cooperativeness; Stakeholder or shareholder value
Intra-organizational aspects of public administration	Robustness; Adaptability; Stability; Reliability; Timeliness; Innovation; Enthusiasm; Risk readiness; Productivity; Effectiveness; Parsimony; Business-like approach; Self-development of employees; Good working environment
Behavior of public-sector employees	Accountability; Professionalism; Honesty; Moral standards; Ethical consciousness; Integrity
Relationship between public administration and the citizens	Legality; Protection of rights of the individual; Equal treatment; Rule of law; Justice; Equity; Reasonableness; Fairness; Professionalism; Dialogue; Responsiveness; User democracy; Citizen involvement; Citizen's self-development; User orientation; Timeliness; Friendliness

In connection with the previous chapter on attitudes and values, Jørgensen and Bozeman attempted to cluster the 72 values according to three different dimensions: proximity, hierarchy, and causality. The first dimension is about the proximity between two (or more) variables to each other. This means that two variables can be either (1)unrelated; (2) neighbor values, (3) co-values, or (4) nodal values. The first category is self-explanatory, with no obvious connection between the two values. Neighbor values, such as "parsimony" and "productivity", are close but not identical. Jørgensen & Bozeman assumed that the number of neighbor values a single value has determines the importance of a value, as the single value is more equipped with different facets(Jørgensen & Bozeman, 2007, p. 370). Co values mean that one value may positively affect another value, or that one value acts as a precondition to another. Lastly, the category of nodal values describes a single value that is related to a large number of other values.

The last two dimensions of public values concern hierarchy and causality. Jørgensen & Bozeman persist that the hierarchy of values depends on what they call "relative primacy"(Jørgensen & Bozeman, 2007, p. 373). This conceptual relationship between values is similar to the assertations by previously mentioned authors such as Ravlin & Meglino (1987). The hierarchy of values is thereby determined via the preferred choice of one value over another (primacy) but is always seen in the context of a comparison between two or more values against each other (relative). Based on this definition of value hierarchy, Jørgensen & Bozeman reiterate that there are two sorts of core-value categories namely prime values and instrumental values. While prime values are defined as "ends in themselves" is the instrumental value seen as something that helps to lead to another, stronger value (Jørgensen & Bozeman, 2007, p. 373).

However, as Rainey (2009) pointed out, the values that public organizations pursue are generally diverse, multiple, and often conflicting with each other (Rainey, 2009, p. 68). Moreover, Bozeman's last characteristic of public value is represented in the last century's different public administration paradigms. The following paragraphs illustrate these paradigms and concentrate on their respective public values, including the causality of prime values and potential anteceding instrumental values, as well as determining neighbor-, co-, and nodal values. Afterward, the thesis will go over to develop specific hypotheses to be tested.

2.3 How Values affect perception and attitudes

What are values rooted in? In the words of Mark Moore: "Value is rooted in the desires and perceptions of individuals – not necessarily in physical transformations and not in abstractions called societies"(Moore, 1997, p. 52). But, in general, how do values translate into perceptions or, as it is framed in the sociology literature, the term "attitudes"? What kind of role do they play?

Generally, sociology academics use two models of values to categorize them: "values as preferences" and "values as principals". The first concerns individuals' preferences towards something values and the resulting attitudes towards (?) their values. The second one deals with values that determine a deeper, personal guideline of how an individual should act (Parks & Guay, 2009). The thesis at hand will deal with the first category.

In the academic sphere of behavioral studies, there are numerous examples of studies that link values with attitudes or behaviors (Ajzen, 1991; Gerard & Orive, 1987). However, this thesis focuses on the formation of perceptions and attitudes. Specifically, how one or more values influence perceptions and attitudes. Regarding the prioritization of values against one another, there are two different conceptualizations of value measurement methods: the first one sees the relationship between values as hierarchical, which implies that individuals rank their preferred values and choose their behaviors and actions accordingly. The second sees values as something that does not necessitate hierarchy, meaning that two or more values can simultaneously be equally important to the individual (Ravlin & Meglino, 1987). According to Ravlin and Meglino, the latter conceptualization offers intersubject comparisons for value preferences between individuals, while the first conceptualization does not (Ravlin & Meglino, 1987, p. 666). However, the findings of their study supported arguments of the first conceptualization that individuals have a hierarchical ordering of their values in their memory, which they apply when they find themselves in uncertain situations(Ravlin & Meglino, 1987, p. 672)

Moreover, Homer and Kahle (1998) suggested a relationship between values and attitudes in line with hierarchically ordered values. In line with previously developed social adaptation theory, Homer and Kahle argued that values are the most basic, but abstract, form of social cognition and can thereby be considered as "prototypes" from which attitudes and behaviors are formed (Homer & Kahle, 1988, p. 638). Both authors tested their proposed pyramidical relationship from values to behaviors in their study. They used a List of values (LOV)::

A sense of belonging (1);

excitement (2);

fun and enjoyment in life (3);

warm relationships with others (4);

self-fulfillment (5);

being well-respected (6);

a sense of accomplishment (7);

security (8)

and self-respect (9).

They used this list to test how these values relate to nutrition attitudes and shopping behaviors. The analysis indicated that there is indeed a relationship between the selected values and nutrition attitudes, while the relationship between the values and shopping behaviors was mediated by the nutritional attitudes.(Homer & Kahle, 1988, p. 647) While the authors acknowledged it to be difficult to apply this to other attitude categories, the study proved another dimension of the value-attitude relationship to be of importance, namely value interconnectedness: The authors distinguished between different dimensions a value has (internal/external and "apersonal"/personal) and that depending on the dimension these values fulfill and the importance that individuals assert to dimension, the degree of association towards nutrition attitudes changes. (Homer & Kahle, 1988, p. 645).

In extension, a more recent study by Van Overwalle and Siebler (2013) tried to combine the different streams of attitude formation and change theories. The model by the authors displays the association of attitudes and values as interactions of nodes within a network where attitudes towards objects are reinforced by experiences about an attribute and the attributes are evaluated based on these experiences ("*valence*"). The valence of the experience reconnects with the attribute when the individual evaluates an object. If the object possesses positive (or at least more positive than negative attributes), the object is also positively evaluated.(Van Overwalle & Siebler, 2013, pp. 235-237). One component of their study included the integration of expectations and values into a model for attitude predictions. In their reiterations about this component, the authors engage with the core ideas behind the theory of the expectancy-value model. According to the authors, the more often an attitude is evaluated based on an attribute, the stronger the connection between the attribute and the valence. This affects the overall attitude towards an object in such a way, that the higher occurrence of one or more attributes that are positively connotated can lead to a generally positive evaluation of the object, regardless of existing attributes with negative connotations that the individual has previously established towards the object (Van Overwalle & Siebler, 2013, p. 240). This means the stronger the cognitive connection between an object and a respective positive attribute, the more likely it is that the object is perceived positively.

Concerning this thesis, the mentioned theories on attitude formation and change offered three important insights: First, research in this field suggests an association between values and attitudes, where values are the foundation to develop attitudes. Secondly, as Ravlin and Meglino suggest, there is merit in investigating the relationship between different values and finding ways to compare different degrees of preference between these values meaningfully.(Ravlin & Meglino, 1987, p. 672). Thirdly, there are suggestions in previous studies that values may have antecedent values that build on each other in a consecutive order.

The next chapter will build upon these assumptions by presenting different kinds of public values and presumed interconnectivity between them. After that, the thesis will present hypotheses that are built upon public reform paradigms and their respective set of public values they represent.

2.4 Administrative paradigms

The next two chapters will with and present two prominent administrative paradigms. A selection of values for each paradigm from the Jørgensen & Bozeman inventory of public values will be presented. However/ Although, the emphasis is laid on "selection", as there are multiple options of public values that could be argued for. Table (2), is an unexhaustive list of values that are shared between the paradigms, as well as values that are exclusive to one of the paradigms.

Table 2

List of shared and exclusive values of the TPA and NPM paradigm

Exclusively TPA value	Shared values	Exclusively NPM value
Loyalty	Reliability	Efficiency
Rule of law	Accountability	Parsimony/Productivity
Expertise	Responsiveness	Performance
Stability		Business-like approach
Neutrality		Risk readiness
		Fairness

This thesis is particularly interested in the public values of "rule of law" and "expertise" concerning TPA values, on the other hand on the public values of "efficiency" and "performance". The next two paragraphs will discuss them in more detail.

2.4.a Traditional Public Administration (TPA)

Starting with the paradigm that represents "traditional" public administration (TPA), one of the biggest inherent public values of this paradigm can be expressed by what Jørgensen & Bozeman described as "expertise": While the political sphere determines the goals, the executive branches in the form of "field experts" act as advisors and implementers to refine and slightly steer to meet the politically determined goals (Bryson, Crosby, & Bloomberg, 2014, p. 446; Rhodes, 2016, p. 2). Traditionally, a civil servant was thus much more regarded as a neutral expert with long-standing experience in their field. This vision of the neutral expert is part of the dichotomous separation within the state, where the political sphere adheres to questions about legitimacy and democratic support, while the civil servant assesses and manages the need for resources to reach the targets set by politicians(Moore, 1997). Specifically, the sense of dichotomy between the administration and the political sphere by civil servants as well as what Lynn (2001) in his work framed as "neutral competence" (Lynn Jr, 2001, p. 146) is one of the core understandings of how public administration used to be regarded as the most successful prototype (Katsamunska, 2012, pp. 75-76). Scholars in the field of public administration such as Woodrow Wilson, or Frank Goodnow extensively wrote about the importance of isolating the administrative tasks and responsibilities of government.

Another public value that is ascribed to the traditional sense of public bureaucracy by Denhardt & Denhardt (2000) is “the rule of law”, where the decisions that are being made are enshrined in a code or a rule book that acts as a safeguard to limit administrative discretion for reasons of legitimacy.

Profile Public Administration: (Denhardt & Denhardt, 2000, p. 551). The adherence to proper procedures and systems was designed to ensure fair and equal treatment to citizens and the ruling government, irrespective of their political affiliation (Katsamunska, 2012, p. 75). “The rule of law” belongs to a public value category of “covalues”, as there are many public values that directly affect the rule of law, such as “openness” or “robustness”, but also “justice”. In this case, “the rule of law” can be seen as a covalue for the public value of “equal treatment” of civil servants towards citizens. This does, however, not include citizens being actively involved in policymaking, as there is the presumption that the citizens look out for their self-interest, rather than the public interest (Denhardt & Denhardt, 2000, p. 552).

While both of these public values were considered important when the TPA was the dominant paradigm, these two values lost their stance in later paradigms. They did not dissolve or drift into obsolescence but were considered less important in contrast to opposing public values. Both public values of “expertise” and the rule of law” are connected to the administrative model of a *Rechtsstaat*, which places the administrative state as an integral part of society, which is concerned with the “[…] preparation, promulgation, and enforcement of laws” (Pollitt & Bouckaert, 2017, p. 61). These essential tasks are executed by civil servants with legal and technical expertise within their field to exsert legal control (Pollitt & Bouckaert, 2017). However, this model starkly contrasts the Anglo-Saxon typical ‘public interest’ model, which was promoted during the NPM paradigm. Essentially, the ‘public interest’ model regards laws as a tool of governance, but the underlying procedures and tangibles are not as dominant as in the *Rechtsstaat* model (Pollitt & Bouckaert, 2017, p. 61). Thereby, civil servants are not trained and skilled in the law under the ‘public interest’ model but are preferred as “generalists” (Pollitt & Bouckaert, 2017). , The public values of “expertise” and “rule of law” becoming less prevalent under the ‘public interest’ model dominated NPM reforms.

Having said that, how do preferences towards these two public values translate into the perception of public sector reform success? Here, the theory by Van Overwalle & Siebler provides theoretical foundations: It was already explained in the previous theory section what the expectancy-value theory entails, where an object is evaluated based on the exchange aggregated amount of positively (or negatively) associated attributes (Van Overwalle & Siebler, 2013, p. 238). Henceforth, positive aggregated opinions towards the public values of “expertise” and “rule of law” should theoretically negatively impact a civil servant's perception regarding the success of reform success:

H1A: The higher the degree a civil servant sympathizes with the public value of “expertise”, the lower they will perceive the degree of reform success.

H1B: The higher the degree a civil servant sympathizes with the public value of the "rule of law", the lower they will perceive the degree of reform success.

2.4.b New Public Management

In light of the presented criticism of TPA, scholars have started to develop methods that adhere to running the government like a business and using private-sector methods and practices since the 1970s. First, and foremost, the countries which represented the dominant forerunners of these public management reforms were OECD countries (Hood, 1995). Even though the term "New Public Management" (NPM) is in itself fairly vague and academically contested (Denhardt & Denhardt, 2000, p. 551; Dunleavy & Hood, 1994, p. 9). Although, there is a certain set of characteristics that are reoccurring in academic literature. While the more radical, mostly Anglo-Saxion countries questioned the basic idea of whether public services are necessary to be delivered by the public sector, the European countries were more focused on renovating the public sector to make use of private sector management provisions to circumvent or solve inefficiencies of the current system (De Vries, 2016, p. 37). However, the degree to which these values and characteristics are present varies within countries' administrative systems, as there is no such thing as one type of NPM administrative regime (Pollitt & Bouckaert, 2017, p. 72).

Hereby, the internalized public values of traditional public administration shifted. Before that, core elements of the TPA included values such as the "rule of law" and "ethos for public service" professionals, which were the pillars of the public service. These pillars were rooted in hierarchies and civil servants' expertise within the field. After the shift, efficient use of resources and effective delivery of services are supposed to be supported by assumptions that have been taken from private sector practices and economically rational market mechanisms such as rational choice-, principal-agent- and transaction cost theories (Hood, 1991). Therefore, a shift away from hierarchies within the public sector towards markets and formal contracts was seen. Due to the out-sourcing of public services to private sector actors, the government became more fragmented, and civil servants experienced less discretion within their field.(Dickinson & business, 2016, p. 43) Moreover, public managers started to develop a more innovative and entrepreneurial mindset regarding the creation and delivery of public goods. Tools and techniques to supposedly encourage this undertaking include scientific and quality management methods (Denhardt & Denhardt, 2000).

These changes are in line with public values that belong to a subgroup of the category "Intra-organizational aspects of public administration" in the inventory by Jørgensen & Bozeman, (2007), namely the subgroup "productivity". In general terms, productivity refers to the amount of output that is provided per unit of input. There have been extensive discussions around the subject of how output is assumed to be valued in the government sector, especially the default to assume that the overall governmental output is equal to the total value of all inputs (Boyle, 2006, p. 4). Scholars such as Pollitt and Bouckaert (2004) came up with different alternatives for how to value output and thereby also how productivity might increase: (1) resources (inputs) decrease and outputs increase; (2) resources remain the same and outputs increase; (3) where resources increase but outputs increase by an even larger amount; (4) outputs remain static but resources decrease; (5) where outputs decrease but inputs decrease by an even larger amount (Pollitt & Bouckaert, 2017, p. 143). However, the problem with probable output measurement does not end there, as it is already acknowledged that it is challenging to measure the productivity of an output that was produced by several inputs (Simpson, 2009). It is thereby difficult to asses the efficient use of inputs that are allocated to the production of outputs. This poses a challenge for civil servants and public officials to determine the proper measures for productivity, which typically goes in hand with the second NPM public value of this thesis, namely "performance".

The second public value or the NPM paradigm is "performance". According to Bozman's "Balance Model", the public value performance is strongly connected to NPM. The public value "performance" refers to the extent objectives are achieved by the administration (Bozeman, 2007, p. 80). This is evident in the shift of responsiveness to policies, as civil servants were held accountable to reach predetermined goals. This approach was framed as the "Managing for results" framework. Even though these goals were set and demanded to be fulfilled by the public and the political sphere, civil servants often lack influence on the determination of these goals (Moynihan, 2006). This is where both the NPM public values cross, as "productivity" is frequently used to measure "performance" in the public sector. One reason for that is that the focus on measuring outputs is easier to monitor than other measures of performance (Stainer & Stainer, 1998). However, the inflationary use of performance and productivity measures was not necessarily welcomed in every regard, as some argue that this process commodifies the inputs and outputs of public institutions that are based on societal merit and goodwill such as schools and hospitals (Hörnell & Hjelm, 1994). Moreover, the focus on limiting inputs and still maximizing outputs thereby enhancing productivity and the overall performance of public institutions was a core concern during the NPM reforms (Pollitt & Bouckaert, 2017, p. 128) In the context of the previous theoretical reiterations by Van Overwalle & Siebler about the aggregated amount of associated attributes, civil servants who welcome this influx of performance and productivity measurements of their tasks may as well welcome the reform process, thereby stating:

H2A: The higher the degree a civil servant sympathizes with the public value of "performance", the higher they will perceive the degree of reform success.

H2B: The higher the degree a civil servant sympathizes with the public value of "productivity", the higher they will perceive the degree of reform success.

Conclusion

In Summary, the paradigm of Weberian administration was defined by the pretense of separating the political sphere from the administrative capacity of a state. Because of that, the administration was defined by neutrality and the aspiration to see administrative tasks as a science and therefore building on expertise within the workforce. By treating everyone within the citizenry as equal, effective delivery of public services is marked by focusing on designing the internal process as efficiently as possible, rather than focusing on the output. To still be as efficient in delivering high-quality output, procedures, and steps are as clear and closely defined and followed as possible.

In contrast to TPA, The New Public Management paradigm tried to erase some of the pitfalls of the predeceasing paradigm. The strong emphasis on using limited available resources as efficiently as possible to reach pre-determined goals by using private sector models such as rational choice is one of the key factors that defines this paradigm. Using pre-determined targets opened a different angle to ensure government accountability. Furthermore, developing a customer-focused attitude to engage with citizens and offer fair access to services gave the pretense to tailor the services to the individual needs of the consumer of public goods in the most efficient way (Stoker, 2006, p. 50).

In the context of the inherent public values for each of these paradigms, the underlying theoretical assumption suggests that civil servants who emphasize TPA values do not perceive the reform in the public sector positively, while civil servants who sympathize with public values that relate to the NPM paradigm perceive the public sector reform more positively. The previously stated hypotheses are summarized in the model (figure 1)below:

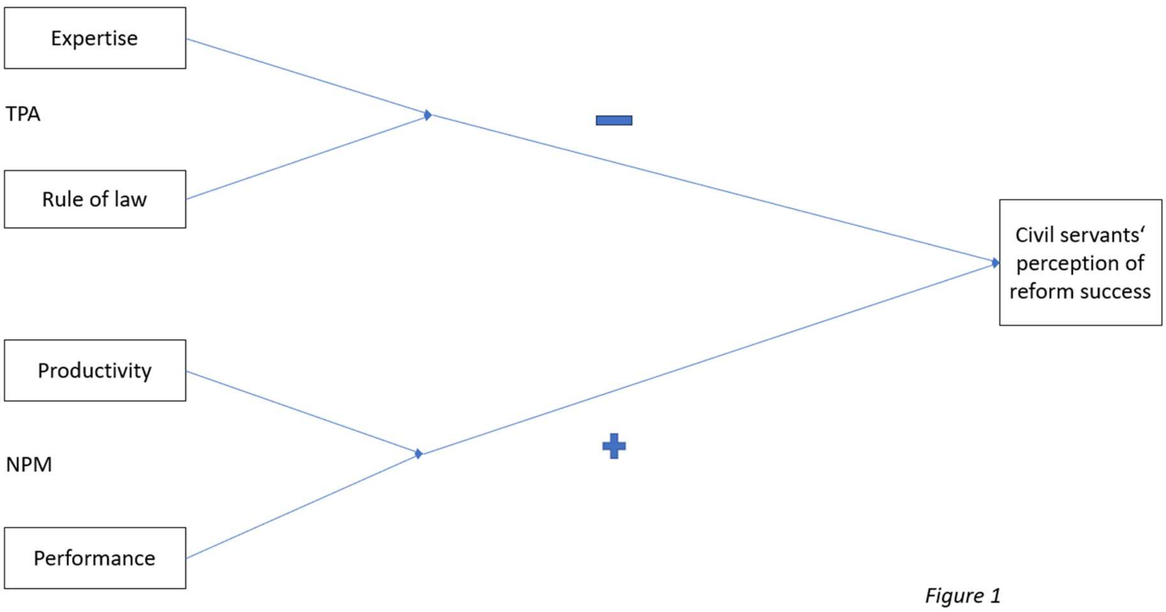


Figure 1

3. Methodology

3.1Research design

This thesis uses a quantitative research design based on secondary data. This method was chosen instead of qualitative research and primary data such as interviews. Using quantitative data in this cross-sectional research design setup enables us to use aggregate data to determine the influence that individual value preferences might have on the perception of public sector reform. Primary data such as interviews can be useful afterward, once a potential relationship between public values and the perception of public sector reform success is established.

3.2 COCOPS Dataset

The COCOPS Dataset is an accumulation of surveys conducted in eight different EU countries, Norway, and the UK. The European countries include Austria; Estonia; France; and Germany. There have already been numerous publications that split the dataset into diverse directions of public management research such as structural and organizational commitment (Suzuki & Hur, 2020), trends and effects of social cohesion in Europe (Andrews, Downe, Guarneros-Meza, Jilke, & Van de Walle, 2013), effects regarding principal-agent theories (Bach, Hammerschmid, & Löffler, 2020) as well as the general effects of NPM reforms in a comparative setting between several countries (Hammerschmid, Van de Walle, Andrews, & Mostafa, 2019; Kickert, Randma-Liiv, & Savi, 2015). Given that national datasets do not always comprehensively consider all institutional levels within a country, the dataset at hand opens up the possibility of gaining a more thorough picture of a country's administration as other researchers have discovered already (Galli, Ongaro, Ferrè, & Longo, 2013). Intra-national comparisons have the advantage of providing insights into potential cultural, regional, or ideological groups and differences within a country that might have influenced the decisions that lead to the present level of public sector reform efforts (Pollitt & Bouckaert, 2017, p. 28).

The Data that will be used in this thesis is derived from the COCOPS survey, which was conducted between 2011 and 2013 as a part of "Work Package 3" of the COCOPS project (Hammerschmid, Oprisor, & Štimac, 2013, p. 4). This project was intended to give insights into the general trends in the public sector by professionals in the field as well as employees in the public sector (Curry, Van de Walle, & Gadellaa, 2014, p. 2). The survey was sent to 20,000 employees from 10 different European countries. Around 4814 respondents completed the survey, meaning that the survey had a response rate of 23.71%.

The second part deals with the management and work practices that are present. The third category goes into the public sector how it has been reformed and how the "fiscal crises" have affected administrative behavior. The last part is primarily concerned with the attitudes, preferences, and personal information of the civil servant. Thus, digging more into the motivational aspects and background circumstances of the respondents.

As the dataset is a merged cluster of several, sometimes independently, conducted surveys on the national level (for example Germany, where the Federal Employment Agency conducted a separate survey version), there are some country-specific variations (Hammerschmid et al., 2013, pp. 12-13). In general, the response rate was regarded as low in contrast to previously conducted surveys in different European countries as well as in the US, which was justified with several constraints such as concerns regarding the anonymity of the survey, the sensitivity of the gathered data and the high workload of the concerned civil servants. In addition to that, the previous observation has been made that survey response rates are generally lower in longitudinal studies (Hammerschmid et al., 2013, p. 30). Regardless of the low response rate of the survey, the dataset holds merit in answering the research question in different ways: First, the dataset contains civil servants' responses from both federal and Länder levels. Thus allowing to put emphasis on the reported discrepancies between both governance levels and to investigate key differences. Secondly, it holds the personal perceptions and motivations of these employees, making it eligible to be used to research the potential motivational shifts that the change of paradigms anticipates.

Even though this data set was conducted nearly ten years ago before today's date, it can still be considered relevant when it comes to the evaluation of public sector reform. The reason for that is that this dataset is the most comprehensive collection of data concerning insights for public sector reform in Europe. There has not been any standardized collection of data concerning public sector reform that covers so many different countries within Europe.

3.3 Case selection

This thesis will select country-specific cases, namely cases from Germany. The reason is to limit potential cultural influences on the results, as it is difficult to assess how much of a role they play. According to the official final research report "COCOPS Work Package 3", the survey includes 566 units from Germany. However, the codebook of the dataset and the dataset, both downloaded via the website of the Leibniz Institute for Social Science (GESIS) reports only 445 respondents for Germany. Both the unit of Observation and the unit of analysis are the surveyed employees of these organizations. This thesis will rely on the Dataset published by GESIS[1] and will therefore work with the general stock of N=445 respondents. The German case of public reform is special in the regard, that Germany not only implemented NPM measures as one of the latest countries in the OECD higher ranking countries, but they also did not import a new system but rather improved the existing administrative system (Pollitt & Bouckaert, 2017). Also, changes to the system were developed within the public sector, rather than taking up external private-sector consultants to create solutions (Hammerschmid & Oprisor, 2016). Thereby the reforms are built up with internal competency and less external or foreign interference. Furthermore, Germany was one of the countries that did not immediately opt to minimize state capacities during the height of NPM reforms, but rather deregulate and streamline their bureaucratic apparatus(Pollitt & Bouckaert, 2017, p. 116). However, concerns regarding a bloated public bureaucracy and voices that call for the closure of agencies have been present in the election campaigns (Kiefer,

[1] https://dbk.gesis.org/dbksearch/SDesc2.asp?DB=E&no=6599, last visit 11.05.2023

2025). Table (3) displays the sample size of German civil servants that were included in the dataset, based on the organizational type they belong to.

Table 3
Organization type

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Mini. at central gov. lvl	58	13.0	13.0	13.0
	Agency/subordinate gov. body central gov. lvl	191	42.9	42.9	56.0
	Min. at state/regional gov. lvl	196	44.0	44.0	100.0
	Total	445	100.0	100.0	

3.4 Operationalization of independent variables connected to TPA

The following paragraphs will operationalize the dependent variables. They will start with the TPA-related values. Afterward, the public values connected to NPM will be operationalized. Later, both the variables that are connected to the TPA paradigm ("expertise" and rule of law") are combined into one mean index. The same is conducted for the variables of the public values that are connected to the NPM paradigm ("productivity" and "performance").

Expertise

"Expertise" in the context of public administration refers to the expert knowledge civil servants possess within their specialized public domain (Katsamunska, 2012, p. 80). The thesis will use statement 5, item 6 to measure the degree to which civil servants value the provision of expertise. The statement says: "I mainly understand my role as a public sector executive as providing expertise and technical knowledge". The scale of the statement goes from "Strongly disagree" (1) to "Strongly agree" (7).

Rule of law

The rule of law refers in the context of public administration to the adherence to "administrative" law, which acts as binding guidance for civil servants to ensure previously mentioned pillars of traditional administrative characteristics such as "neutrality" or "stability" (Lynn Jr, 2001, p. 146). At its peak, adherence to the rule of law can lead to ignorance of the outcomes that the administration produces (Gaebler, 1993). This paper uses this extreme to measure the degree to which civil servants emphasized the rule of law by utilizing a recoded version of question 23, item 3 of the dataset: "Public services often need to balance different priorities. Where would you place your position?". The item is designed as a dummy variable coded as "0" with the label "Not following rules", while option 1 is labeled as "following rules" and stays the same[2].

[2] The original item had a different measurement scale and was designed as a trade-off and entailed the scale from "Following rules" (1) on the one end and "Achieving results" (7) on the other end. The re-design into a dummy variable was to ensure a coherent variable design as "achieving results" is not considered to be the same as "not following rules".

3.5 Operationalization of independent variables connected to NPM

Productivity

The measurement of productivity in the public sector context is predominantly focused on the accurate measure of outputs a public organization produces. However, it is more accurate to measure the aggregate amount of inputs (resources) against the aggregate outputs (delivered services) by the organization (Simpson, 2009, p. 251). The measure of the public value "productivity" is focused on outputs and outcomes and thereby uses item 6, from question 8 of the dataset, which states: "To what extent do the following statements apply to your organization?" and item 6 "We mainly measure outputs and outcomes". The scale goes from "Strongly disagree" (1) to "Strongly agree" (7). In terms of concept validity, it is noted that the concepts of outputs and outcomes are inherently different from each other, as output focuses on the services and activities a public organization provides, while outcomes focus on the potential effects to which the output potentially contributes. Moreover, there are voiced concerns about including outcomes in the production measurement of public institutions, as they might not have control over changes (Boyle, 2006, p. 5).

Performance

"Performance" in the context of public administration is commonly addressed by the issue of determining appropriate measurements of the actual performance a public organization produces (Parhizgari & Gilbert, 2004). However, this thesis is not interested in the different measurement approaches, but in how a civil servant values the "performance" of a public organization. Therefore, statement 9 of the dataset will be used, including all 8 items combined into an index. The statement says: "In my work, I use performance indicators to […]" (1) "Assess whether I reach my targets"; (2) "Monitor the performance of my colleagues"; (3) "Identify problems that need attention"; (4) "Foster learning and improvement"; (5) "Satisfy requirements of my line manager"; (6) "Communicate what my organization does to citizens and service users"; (7) "Engage with external stakeholders (e.g. interest groups)"; (8) "Manage the image of my organization". The scale starts with "Not at all" (1) and ends with "to a large extent" (7). In this measure, high overall usage of performance indicators would translate into a high emphasis on the public value of "performance".

3.6 Operationalization of the dependent variable

The "perceived success" of public sector reforms

In measuring the success of a policy, specifically the reform of policies, researchers in the field of public administration can considerably deviate from one another. One reason is that there were only attempts to design a common framework to measure policy success. (Newman, 2014), and none of the frameworks have become commonly accepted. Newman stresses in his academic paper that due to the nature of success and failure being "inherently subjective concepts", it would be rather difficult to establish a framework that encompasses all facts of subjectivity. (Newman, 2014, p. 192). This thesis will therefore use question 18 of the COCOPS dataset, which simply states: "Please indicate your views on public sector reform using the scales below. Public sector reforms in my policy area tend to be […]". There are 11 items for this question, of which the 10^{th} will be used for this thesis as it is the one that measures the perception of success in the most direct way. It measures from "Unsuccessful" (1) to "Successful" (10).

Using this item has different advantages and downsides to it. One advantage is that is very broad and open and, therefore, gives the respondent the chance to internalize its own weighs and criteria rather than adhering to a given set that may not represent the respondent's idea of what encompasses success. However, a major disadvantage is that there is no transparent idea of what criteria the respondent uses. Secondly, it does not give the respondent the chance to give a detailed

message of what they might see as successful within their view and what counts as a failure. The table below table (4) presents all relevant variables. All variables are measured as Likert scales and treated as metric scales.

Table 4

Table of variables

Variable	Relevant Hypothesis	Measured	Item label
Expertise	1	Likert scale (1-7)	1= "Strongly disagree" 7= "Strongly agree"
Rule of law	1	Likert scale (1-7)	1="Following rules" 0= "Not following rules"
Productivity	2	Likert scale (1-7)	1= "Strongly disagree" 7= "Strongly agree"
Performance	2	Likert scale (1-7)	1= "Not at all" 7= "To a large extent"
Reform success	Dependant variable	Likert scale (1-10)	1= "Unsuccessful" 10= "Successful"

3.7 Validity and Reliability

The dataset was derived from the EU COCOPS initiative as part of the third work package, created by five different teams from European academic institutions including the Hertie School of Governance Berlin, the National Center for Scientific Research, the University Pantheon-Assas Paris 2, Cardiff University, Erasmus University Rotterdam and the University of Bergen(Hammerschmid, 2015, p. 4).

To assess the power of the sample and the corresponding effect size, an F-test via the program G-Power 3.0 was conducted, as is seen in appendix 1. The most difficult part of calculating the minimum sample size and the power of the test was to find an appropriate effect size. According to Cohen (1992), the effect size (f^2) can be calculated by using the R^2 of an associated study with the same dependent variable and dividing it by 1-R^2 (Cohen, 2016, p. 157). This thesis therefore uses the R^2 from Nayernia et al (2020) with R=0,474, meaning R^2=0,225. The final result would mean that $f^2 = 0.2$ which will be used for the test. Other than that, the error probability for a Type 1 error (false positive) is set at $\alpha = 0.01$ and the error probability for a Type 2 error (false negative) is set at 0.95. the number of predictors, based on the variables that are used in the model is set at 6. The result is portrayed in the appendix and determines that the sample size for the linear multiple regression needs to be at least 144 respondents.

Internal validity

Concept validity

Concerning the congruency of items that are used for indexes, a Cronbach's Alpha test was performed on the items on the "performance" index. The test gives a number between 0 and 1 to determine how consistent the items of an index are and how strongly these items are related to each other. The test gives also an estimate regarding the possibility of an error in the measurement at hand. (Tavakol & Dennick, 2011)

It is advised that a good Cronbach's Alpha result varies between 0,6, as below would indicate that the items measure content that varies too much from another, and on the top 0,8, as higher would question how useful the index is as the measures of the items would be too close to each other. This would mean that more than one item would measure the same things within an index making the item redundant (Tavakol & Dennick, 2011).

Indices

Table A1 in the appendix shows the results of Cronbach's alpha test for the index on "productivity". The Cronbach's alpha value of 0.938 indicates that the items are too closely related to each other. This is also reflected in Table A1 in the appendix where the value stays above 0.9 regardless of the deletion of items. In this case, Tavakol & Dennick (2011) suggest reducing the test length by reducing the number of items. However, a meaningful reduction that leaves the index with a few items was not possible in the case of the index for "performance". It is therefore decided that the index will be resolved and instead, only item 1, question 9 of the COCOPS dataset ("In my work, I use performance indicators to […]" (1) "Assess whether I reach my targets") will be used as a measurement for the variable "performance". Concerning Cronbach's alpha value for the mean index of the combined TPA public values, the output shows a value of 0.53. This would mean too-high differences in variances from at least one of the included items. One potential explanation for that could be the recoded variable for the "rule of law", which changed from a 7-point Likert scale into a dichotomous variable. For the index on the combined public values "productivity" and "performance" for the NPM paradigm, Cronbach's alpha value is at 0.66. This value is still within an optimal range, as it is between 0.6 and 0.8.

Measurement validity

Lastly, as there are many predictor variables, a test for multicollinearity is conducted. Multicollinearity would mean that there would be a nearly perfect linear relationship between explanatory variables (Alin, 2010). It may be the case that some independent variables may show a statistical association, however, certain levels are tolerable. This is why a test for collinearity was conducted. The output of the test is presented in Table A3 in the appendix. A tolerance value of below 0.1 as well as a VIF value of 10 would indicate multicollinearity between the dependent variables (Daoud, 2017, p. 5). However, none of the variables have a VIF value of 10 nor do they have a tolerance value below 1.0. Based on these test results, the possibility of having multicollinearity between the dependent variables of this model is fairly low.

Regression diagnostics

First, it needs to be tested whether the items are normally distributed, which would be the basis for running regressions. Both values for skewness and kurtosis are displayed in the appendix table A4. Only the recoded variable on the "rule of law" will not be compared, as it is a dichotomous variable. Concerning skewness, all variables are normally distributed, except "expertise". The variable is negatively skewed with a value of -1.365 (expertise) as portrayed in the appendix in their respective histograms. This is reflected in the descriptive data above where all means are smaller than their respective mode (except the dependent variable). Regarding tendencies for the kurtosis, the outputs vary throughout the different variables. All variables, besides "expertise", show a negative kurtosis, as the peak of the observed value is lower than the respective value of a normal distribution.

Table 5
Tests of Normality

	Kolmogorov-Smirnova			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Expertise	.250	399	<.001	.804	399	<.001
Rule of law	.541	399	<.001	.186	399	<.001
Productivity	.145	399	<.001	.941	399	<.001
Performance	.131	399	<.001	.938	399	<.001

a. Lilliefors Significance Correction

The tests conducted for the 4 independent variables are portrayed in table (5) where both the Kolomgorov-Smirnov test and the Shapiro-Wilk test were performed to statistically evaluate the normality of the sample. A significant result (<0,01) in one of these two tests means that the model does not fit the predicted values and the value of the sample is not normally distributed. The present results can be characterized as daunting because none of the variables can be considered normally distributed, according to both test results.
However, there are high doubts about using these tests for samples larger> than 200. The Kolmogorov-Smirnov test is considered to be biased in rejecting normal distribution, as the probability of rejecting normality gets higher with the enlargement of the sample size (Lazariv & Lehmann, 2018). The most reliable method to assess the sample's normal distribution is by analyzing the Q-Q Plots for each variable. All normal Q-Q plots have been included in the appendix.
A look at the normal Q-Q plots shows the placement of observed values that are fairly close to the line of the expected values. The exception to this is the variables for the public value of "expertise". All variables show a compressed curve. It can thereby be assumed that based on the graphical analysis concerning the normality of the variables most of them show fairly normally distributed results besides the mentioned exception of "expertise".The consequence is that for the non-normally distributed variable extra attention has to be paid to the potential implications of the results by the regressional analysis.

4. Analysis and results

The analysis starts with the descriptive statistics of the variables as well as the respective regression diagnostics. The upcoming part will first give some basic descriptives of the stated variables. It will then go on to examine the two stated hypotheses. All tests were performed with a two-tailed alpha of <0.05.

4.1 Descriptives

Table 6

Descriptive statistics

	Minimum	Maximum	Mean	Std. Deviation	Variance
Expertise	1	7	5.96	1.170	1.368
Rule of law	0	1	.0362	.187	0.35
Productivity	1	7	3.87	1.612	2.597
Performance	1	7	3.80	1.790	3.205
TPA index	1	7	3.17	1.058	1.120
NPM index	1	7	3,81	1.485	2.206
Reform success	1	10	5.88	2.112	4.461

First, the descriptive statistics in terms of frequencies are briefly examined and presented in Table (6). Concerning the presented means, leaving the dependent variable out of comparison seems rational, as the scale of the variable differs from the one of the independent variables. Comparing the 7-point Likert scales for the rest of the variables, the most interesting values concerning the mean seem to be the variable connected to TPA public values, namely "expertise". It has a mean value above five, meaning that the analyzed respondents tend to be more favorable, at least compared to the other variables, with these public values on average.

4.2 Regression analysis

A full model of the OLS regression analysis was conducted to provide information on how our independent variables' respective models explain the dependent variable's variance. These models include only one of the respective variables or indices, thereby it is not necessary to defer to the adjusted R Square. The respective R Square values are represented in table (7):

Table 7

Descriptive statistics

Variable (model)	R Square value
Expertise(1)	.001
Rule of law(2)	.015
Productivity(3)	.122
Performance (4)	.098
TPA index(5)	.000
NPM index(6)	.145

The table shows that all the models barely explain the variance of the dependent variable, with the highest score on the side of TPA values being the variable for the rule of law at 1,5%. The overall highest R-Square value is provided by the model including the NPM index with 14,5%.

Table 8

Regression Table

Effect	1	2	3	4	5	6
1. Expertise	-.051(.581)					
2. Rule of law		-1.595(.007)				
3. Productivity			.459(<.001)			
4. Performance				.373(<.001)		
5. TPA index					.002(.988)	
6. NPM index						.547(<.001)

Note. Coefficient (p-value)

Dependant Variable: Reform success

The variables that are connected to NPM show significant levels with values of <.001 (Table 8). Therefore, the models can be marked as significant predictors for the perceived success of public sector reform. However, looking at Table (8) for the significance of each variable, it becomes obvious that the data only provide significant levels for the NPM index and the respective public values, which include the variables “productivity” and “performance” The statistics output suggests that there is a significant, positive relationship between the NPM affiliated public values of “productivity” and “performance” with the degree to which a civil servant regards the public sector reform as a success. The analysis output does not support any significant relationship between the independent variable of “expertise” and the dependent variable. However, the variable on the “rule of law” has statistically significant results as the p-value is still below .05 and thereby below the threshold. The models thereby provide arguments for the instance that the perception of the public sector reform success rises by .373 for every point a respondent put on the scale if they mainly measure their outputs and outcomes and.002 points for every point they put in the scale if they use performance indicators to measure their targets. On the other hand, respondents who adhere to “following rules” score -1.595 points lower on their perception of the reform success in the public sector in contrast to respondents who adhere to the category of “Not following rules”.

Using several different models is not practical to create one single regression equation as the estimate of the constant changes throughout the models

.

Table 9

Final results of Hypotheses

Number of the Hypothesis	Hypothesis	Rejected/Approved
H1A	*The higher the degree a civil servant sympathizes with the public value of "expertise", the lower they will perceive the degree of reform success.*	Rejected
H1B	*The higher the degree a civil servant sympathizes with the public value of the "rule of law", the lower they will perceive the degree of reform success.*	Approved/Data support hypothesis
H2A	*The higher the degree a civil servant sympathizes with the public value of "performance", the higher they will perceive the degree of reform success.*	Approved/Data support hypothesis
H2B	*The higher the degree a civil servant sympathizes with the public value of "productivity", the higher they will perceive the degree of reform success.*	Approved/Data support hypothesis

5. Discussion

The results support the claim that there is a connection between public values and perceptions of the success of public sector reform. While the combined mean index for the NPM-affiliated public values of “productivity” and “performance” showed a statistically significant result in the regression analysis, the other index on the TPA-affiliated public values of “expertise” and “rule of law” did not produce any significant results. However, the model that only included the variable on the “rule of law “ produced significant results.

Before diving into the potential implications of what each of the models implies for the respective public values and the following consequences it is worth looking at the broader picture of what the results mean for existing academic works on theoretical aspects concerning public reform evaluation. This thesis provides new input for further research in the line of argumentative policy evaluation, as the previously conducted analysis has shown results that provide new perspectives on potential factors that influence the evaluation of public sector reforms on an individual level. This is novel as previous models for reform evaluation such as the “competing values model” by Boyne (2003) were focused on the organizational level and how they weigh certain values that affect reform evaluation (G. A. J. P. a. Boyne, 2003, p. 220). The same applies to the existing “multiple constituency model” which advocates for the aggregated criteria that stakeholders individually weigh based on their preferences. Furthermore, this thesis uses the concept of public values, which makes it more accessible on an individual level to determine individual preferences. In this regard, Boyne (2003) is fairly vague about how “stakeholders” in the reform process develop mentioned preferences and thereby evaluate respective reform. The results of this thesis thereby help to determine which public values are in trend with current civil servants. Further, it identifies what these preferred versions of successful reform stand for in terms of internalized public values. However, these versions deviate from each other and depend on the respective paradigm that the public values belong to.
Starting with the NPM-affiliated public value of “performance”, the topic of implementing performance indicators (PIs) and their effects have been researched thoroughly in different contexts, whether through the use of performance management systems and potential unintended consequences (Franco-Santos & Otley, 2018), or how the proper usage of performance indicators may lead to a reduction of the possibility of causing goal displacement (Han & Wang, 2020).

The results of this thesis underline that there is general approval for using/applying performance indicators to measure the output of public sector organizations. Where the results of this thesis might add value is the discussion of where these PIs are useful and which PIs ought to be selected (Markic, 2014). In this thesis, the variable “performance” was measured by the respondent's reported usage of PIs for their targets. This could have implications for issues such as job satisfaction or issues concerned with Public Sector Motivation (PSM), as there is ongoing “withdrawal behavior” in connection to the disconnection of civil servants and their tasks (Court & Appiah, 2024). Voluntarily measuring targets that are connected to the individual task may not only give rise to self-efficacy but also curb the need for issues related to PIs such as managerial gaming (Boyne, 2010). In connection to the association between the second NPM index public value "productivity" and perceived public sector reform success, the results mirror the existing reports only partially. It still begs the question of why both input and output/outcome measurements are not significantly associated with the dependent variable. One potential reason is that output and outcome measurement is more positive for intrinsic motivational issues such as goal attainment and thereby strengthens self-efficacy (Gao, 2015). However, the pre-condition for such a transition is an existing "learning" culture, where outputs are used for knowledge management and thereby steer future actions properly (Wescott & Jones, 2007). Another issue that includes using individual PIs is concerned with the ongoing debate around talent management (TM) in the public sector. In this case, properly designed individual PIs can contribute to the TM in the public sector, which can be an important catalyst for general staff commitment as well as the execution of other NPM initiatives (Dougherty & Van Gelder, 2015).

However, the index of TPA values did not yield the expected results. Especially in the case of the public value of "expertise", the results seem somewhat of a surprise in the broader context of contemporary academic literature. The instrumental need for experts in the public sector touches on issues such as 'interactive governance', where civil servants are necessary as facilitators for input by citizens as well as embedding this input based on their expertise in their field (Blijleven, 2023). Overall, the image of "expertise" in the government has been strongly scrutinized in recent years, specifically in connection to the health policies under COVID-19 and yet scholars in the field of public administration are increasingly advocating for more expertise in public bureaucracies (Christensen, 2024). The need for experts in public bureaucracies combined with the statistical output of this thesis portrays a picture of contrasts where experts are needed from a theoretical standpoint, but civil servants may not see the need to be one in the context of a successful public sector reform. Future research may therefore try to approach this contrarious issue and figure out, why civil servants might not feel the need to have a high set of expertise in their field to contribute to a successful public sector reform.

Concerning the results for the public value "rule of law", the results show that against the premises that the rule of law is exchanged for what the authors of the COCOPS questionnaire farmed as "achieving results", "rule of law" seems to be not as vital for public sector reform. Moreover, the public value still has relevance in current issues such as 'Good Governance", where the rule of law is one of the main pillars that provide assurances to involved stakeholders that policy decisions are based on legal principles and thereby promote the idea that the policy decisions are impartial in the context of societal premises (Sari, 2023). However, this issue might soon become more prominent as there are current political movements in the U.S. actively trying to circumvent long-standing processes to change the public sector more rapidly under the guidance of a parsimonious state, as can be observed in the U.S. regarding 'Dodge' department lead by Elon Musk (Fitzgerlad & Honderich, 2025).

Overall, the results do to some extent concur with contemporary academic discussions. However, the merging of two public values into one index does not provide context for which value is more dominant, or which public value might supersede another. This circumstance reveals the disadvantages of the developed theoretical model of this thesis and concurs with problems of previous research: In the paper of Ravlin & Meglino (1987), the authors already suggest that value preferences and order have to be regarded in a network of values, which need other measures than the Likert-scale method. Only rank-order measures that directly compare values reveal the potential cognitive ranking of values and their respective associations. This means that the overall coherency of the results is questionable as one essential part of research between values and attitude formation is lacking, which is the interconnectedness between the public values.

Still, there are practical implications of the analysis that open up potential ways to answer the research question "To what extent do public values explain differences in the perception of German civil servants regarding the success of reform in the administrative system?". The data contributes a clearer understanding of values that civil servants prefer and provides potential answers to which values might contribute to the perception of the success of public sector reforms. The results also reveal the need for a discussion on whether classical public values that are connected to the traditional sense of PA are still regarded as important as values that became more prominent with the NPM paradigm. In line with research, for example, as conducted by Boyd and Nowell (2023), this thesis provided input on how unique concepts of the public sector such as PSM or public values affect practical issues such as the organizational commitment by civil servants (Boyd & Nowell, 2023).

5.1 Limitations of this thesis

The limitations of this thesis can be divided into two specific categories: First, specifically, the measures of public values need better operationalization. The measures for some public values are vague and sometimes arguably misleading. For example the measurements for "performance" and especially for the "rule of law". In the case of "performance", the respondent was asked whether he uses performance indicators to measure the attainability of his targets, while the common definition of performance in the context of public sector performance refers to the degree a public organization delivers goods and services effectively. (Putu S, Jan van Helden, & Tillema, 2007, p. 2). In the case of the "rule of law", the recording into a dummy variable took away potential nuances that the respondents may have wanted to express.

Another reason that limits the degree to which these factors can be tested against the degree of perceived success is the availability of data. The amount of available data to adequately research and test associations between established values of a paradigm to the perception of reforms is rather scarce. Not only was it difficult to find a dataset with a large enough sample size to give at least some sort of prediction for the German Administrative system, but it was also difficult to collect meaningful variables from the dataset to test the created hypotheses. For example, the dependent variable is missing a time constraint. The question in the questionnaire was intentionally posed in a general way, to get information on the general impression of the success of any reform. It is therefore not clear which reforms are meant and in which context these reforms were successful or not. It would have been interesting to create a cross-sectional study at different periods such as in the middle of the 2000s, were the beginning or middle of the implementation of new reforms in Germany was in full process.

Another backdrop of this thesis is the exclusion of 'shared' values, meaning public values that are both present in the NPM as well as the TPA paradigms. One of the values that this thesis is missing however is about "fairness", which was included in a list of NPM values in the "Balance model" by Bozeman (2000). Similar to "neutrality", "fairness" has always been seen in the context of the political *Zeitgeist* as one interpretation leans towards equal treatment of all citizens, while another stresses the preferred treatment for the underprivileged or minorities (Bozeman, 2007, p. 80) Concerning the public value of "fairness" and the suggested results of its significant relation to "reform success", it is hard to find equal support in other studies for the claim. There are studies such as Dean (2021) that connect fairness and anti-public sector bias concerning the performance of the government. (Davis, 2021). Moreover, a better fit for this thesis seems the work by Raaphorst & Van de Walle (2020) about signal theory and communication between citizens and civil servants. In this context, fairness in the process is seen as vital for the citizen's trust in judgment, while on the other hand, the reciprocal fairness from the citizens enables the civil servant to build trust in the honesty of the citizen.(Raaphorst & Van de Walle, 2020). It is thereby difficult to discern facets that belong exclusively to the public value of "fairness", while they do not have any connection to the public value of "neutrality".

Another limitation that could be an interesting idea for future research in this field could be to use a different sample group, namely civil servants from lower governance levels such as the municipal level. Civil servants might have different levels of esteem for the importance of the analyzed public values. The research that is done in this thesis only analyzed the responses of civil servants from ministries or at least national subordinate government bodies on a national level.

5.2 Outlook for further research and recommendations

Concerning the outlook for further research, three things could be addressed. For once, a more coherent approach to include more public values than just two for each of the paradigms. Additionally, as already mentioned, the scope of potential values can be enhanced by including public values that are shared across paradigms. In the context of paradigms, the inclusion of more paradigms can also be an option to deepen the understanding of public values and civil servants' perceptions. Even though NPM and TPA are considered the most prominent ones, several countries developed their respective synthesis of a reform paradigm for the public sector (Pollitt & Bouckaert, 2017, p. 63). Moreover, a clearer differentiation between the values of the public management paradigms could also help in gaining insightful knowledge of the connection between beliefs/values and perception. Here, the paradigms need to be conceptualized based on specific and distinct values that are not shared with other paradigms. However, it is at least questionable whether the use of the public management paradigms is the most promising approach to test the relationship between values and perception. Still, addressing the probably most pressing matter to develop these different approaches and research designs, more data is needed. Specifically, more data is based on questionnaires that are tailored to this particular topic.

Nevertheless, this seems highly improbable, as the developers of the COCOPS Dataset were already faced with the problem of having to delete the last six items, which were deemed too sensitive and had low response rates (Hammerschmid et al., 2013, p. 11). Another approach could include direct interviews with a selection of suitable civil servants. This approach seems more adequate as it is more likely to generate an environment of trust, in which interviewees are more likely to answer such intimate questions. Having a more solid database from which meaningful data stem could help to further refine already existing theoretical concepts as well as develop new ones. The results of this thesis reveal the justification for a more thorough investigation of the tested public values by conducting interviews with specific groups of civil servants. This approach might help to reveal how these results have to be interpreted in the light that somehow past reforms erode or at least threaten the public value of the “rule of law”.

Concerning potential implications for current administrative practices, the thesis provides arguments to look beyond the focus on external stakeholders to identify their position towards public sector reforms and include their workforce more actively in the reform evaluation process as well as the planning on which aspects ought to be reformed. Even though it might be regarded as practical to base reforms on single individuals, it might be useful to include the preferences of civil servants to enhance factors that positively impact the successful execution of reforms such as organizational commitment.

References:

Abdullayeva, N. J. M. н. (2020). ISSUES OF REFORMING ACCOUNTING AND REPORTING IN THE PUBLIC SECTOR. (4 (37)), 6-9.

Ajzen, I. (1991). The theory of planned behavior. *Journal of Organizational Behavior and Human Decision Processes, 50*(2), 179-211.

Alghizzawi, M. A., Masruki, R. J. J. o. M. A., & Auditing. (2020). Issues on Public Sector Accounting Reform in Jordan. *16*(3), 130-139.

Alin, A. (2010). Multicollinearity. *Wiley interdisciplinary reviews: computational statistics, 2*(3), 370-374.

Alves, A. C. (2021). Public Sector Reform in Brazil: changing human capital policies.

Andrews, R., Downe, J., Guarneros-Meza, V., Jilke, S., & Van de Walle, S. (2013). Public management reforms and emerging trends and effects on social cohesion in Europe.

Bach, T., Hammerschmid, G., & Löffler, L. (2020). More delegation, more political control? Politicization of senior-level appointments in 18 European countries. *Public Policy and Administration, 35*(1), 3-23.

Bekelcha, K. L. (2023). Heerrarsa as Instrument of Fighting against Injustice and Oppression in Ethiopia: Notes on

Hachalu Hundessa's Performance on Three Concerts. *International Journal of Leadership and Public Sector Reform (IJLPSR), 1*(2), 166-180.

Bertelli, A. M., & M. Riccucci, N. J. P. a. r. (2022). What is behavioral public administration good for? *, 82*(1), 179-183.

Bhanot, S. P., & Linos, E. J. P. A. R. (2020). Behavioral public administration: Past, present, and future. *80*(1), 168-171.

Blijleven, W. J. L. g. s. (2023). Expert, bureaucrat, facilitator: the role of expert public servants in interactive governance. *49*(4), 841-860.

Bovens, M. (2018). Revisiting the study of policy failures. In *Fiascos in Public Policy and Foreign Policy* (pp. 23-36): Routledge.

Bovens, M., Hart, P. t., & Kuipers, S. (2006). The politics of policy evaluation. In *The Oxford handbook of public policy*.

Boyd, N., & Nowell, B. (2023). Sense of community, sense of community responsibility, organizational commitment and identification, and public service motivation: a simultaneous test of affective states on employee well-being and engagement in a public service work context. In *Public Service Motivation* (pp. 76-102): Routledge.

Boyle, R. (2006). *Measuring public sector productivity: lessons from international experience* (Vol. 35): Institute of Public Administration.

Boyne, G. A. (2003). What is public service improvement? *Public administration, 81*(2), 211-227.

Boyne, G. A. (2010). Performance management: Does it work. In *Public management and performance: Research directions* (pp. 207-226).

Boyne, G. A. J. P. a. (2003). What is public service improvement? *, 81*(2), 211-227.

Bozeman, B. (2007). *Public values and public interest: Counterbalancing economic individualism*: Georgetown University Press.

Bryson, J. M., Crosby, B. C., & Bloomberg, L. (2014). Public value governance: Moving beyond traditional public administration and the new public management. In: Wiley Online Library.

Christensen, J. (2024). When bureaucratic expertise comes under attack. *Public Administration, 102*(1), 79-94.

Cohen, J. (2016). A power primer.

Collington, R. J. N. P. E. (2022). Disrupting the welfare state? Digitalisation and the retrenchment of public sector capacity. *27*(2), 312-328.

Court, T. O., & Appiah, A. K. (2024). Pay dissatisfaction and withdrawal behaviour of employees in the civil service: does personal income tax regime matter? *African Journal of Economic and Management Studies, 15*(4), 654-669.

Curry, D., Van de Walle, S., & Gadellaa, S. (2014). Public Administration as an academic disicpline: Trends and changes in the COCOPS academic survey of European Public Administration scholars.

Daoud, J. I. (2017). *Multicollinearity and regression analysis.* Paper presented at the Journal of Physics: Conference Series.

Davis, J. A. (2021). *Fairness, Performance and the Public Sector: The Role of Fairness in Anti-Public Sector Bias*: American University.

De Vries, M. S. (2016). *Understanding public administration*: Macmillan International Higher Education.

Denhardt, R. B., & Denhardt, J. V. (2000). The new public service: Serving rather than steering. *Public administration review, 60*(6), 549-559.

Dickinson, H. J. T. T. S. S. D. p. p. i. c. w. n.-f.-p., & business. (2016). From new public management to new public governance: The implications for a 'new public service'. *41*, 41-60.

Dougherty, G. W., & Van Gelder, M. J. R. o. P. P. A. (2015). Public agency hiring, minimum qualifications, and experience. *35*(2), 169-192.

Feroz, N. (Producer). (2023, 03.08.2024). 'Eye rolling' civil service leaders block reform, report says. Retrieved from https://www.bbc.com/news/uk-politics-66396350

Fitzgerlad, J., & Honderich, H. (2025). What is Doge and why is Musk cutting so many jobs? *BBC News, London and Washington DC*. Retrieved from https://www.bbc.com/news/articles/c23vkd57471o

Franco-Santos, M., & Otley, D. (2018). Reviewing and theorizing the unintended consequences of performance management systems. *International Journal of Management Review, 20*(3), 696-730.

Gaebler, T. (1993). *Reinventing government: How the entrepreneurial spirit is transforming the public sector*: Plume.

Galli, D., Ongaro, E., Ferrè, F., & Longo, F. (2013). Public sector reform in Italy: views and experiences from senior executives.

Gao, J. (2015). Performance measurement and management in the public sector: Some lessons from research evidence. *Public Administration and Development, 35*(2), 86-96.

Gerard, H. B., & Orive, R. (1987). The dynamics of opinion formation. In *Advances in experimental social psychology* (Vol. 20, pp. 171-202): Elsevier.

Hammerschmid, G. (2015). COCOPS Executive Survey on Public Sector Reform in Europe – Views and Experiences from Senior Executives - reduzierte Version. In: GESIS Datenarchiv, Köln. ZA6599 Datenfile Version 1.0.3, https://doi.org/10.4232/1.12475.

Hammerschmid, G., & Oprisor, A. (2016). German public administration: incremental reform and a difficult terrain for management ideas and instruments. In *Public Administration Reforms in Europe* (pp. 63-72): Edward Elgar Publishing.

Hammerschmid, G., Oprisor, A., & Štimac, V. (2013). COCOPS Executive Survey on Public Sector Reform in Europe.

Hammerschmid, G., Van de Walle, S., Andrews, R., & Mostafa, A. M. S. (2019). New public management reforms in Europe and their effects: Findings from a 20-country top executive survey. *International Review of Administrative Sciences, 85*(3), 399-418.

Han, X., & Wang, W. (2020). *Can Managerial Empowerment Mitigate Goal Displacement?* Paper presented at the 2020 APPAM Fall Research Conference.

Homer, P. M., & Kahle, L. R. J. J. o. P. (1988). A structural equation test of the value-attitude-behavior hierarchy. *54*(4), 638.

Hood, C. (1991). A public management for all seasons? *Public administration, 69*(1), 3-19.

Hood, C. (1995). The "new public management" in the 1980s: Variations on a theme. *Accounting, organizations and society, 20*(2-3), 93-109.

Hörnell, E., & Hjelm, P. J. (1994). Achieving service productivity: lessons from the best in the world.

Huijbregts, R., George, B., & Bekkers, V. (2022). Public values assessment as a practice: integration of evidence and research agenda. *Public Management review, 24*(6), 840-859.

Ishihara, T. J. P. M. R. (2022). Public sector reform and public management theory—cases of Japan. *24*(11), 1653-1662.

Jørgensen, T. B., & Bozeman, B. (2007). Public values: An inventory. *Administration & society, 39*(3), 354-381.

Katsamunska, P. (2012). Classical and modern approaches to public administration. *Economic alternatives, 1*, 74-81.

Kaufmann, L. (2024, 13.06.2024). Kompromisse bei Digitalisierung, Schiene und Straße. *Tagesschau.*

Kickert, W. J., Randma-Liiv, T., & Savi, R. (2015). Politics of fiscal consolidation in Europe: a comparative analysis. *International Review of Administrative Sciences, 81*(3), 562-584.

Kiefer, F. (2025, 11.02.2025). „Möglicherweise ist noch mehr drin": Lindner will vier Ministerien abschaffen und massiv Stellen abbauen. *Tagesspiegel*. Retrieved from https://www.tagesspiegel.de/politik/moglicherweise-ist-noch-mehr-drin-lindner-will-vier-ministerien-abschaffen-und-massiv-stellen-abbauen-13185268.html

Lentner, C., Molnár, P., & Nagy, V. J. E. A.-X. E. Č.-X. (2020). Accrual accounting and public finance reforms in Hungary: the study of application in the public sector. *183*.

Lynn Jr, L. E. (2001). The myth of the bureaucratic paradigm: What traditional public administration really stood for. *Public Administration Review, 61*(2), 144-160.

M'Mugambi, D. K., Okeyo, W., & Muthoka, M. J. A. J. o. E. I. (2020). Public Sector Reforms in Compliance With Public Service Ethics: the Case of County Governments in Kenya. *2*(9), 76-98.

Markic, D. (2014). A review on the use of performance indicators in the public sector. *Tem Journal, 3*(1), 22-28.

Marsh, D., & McConnell, A. J. P. a. (2010). Towards a framework for establishing policy success. *88*(2), 564-583.

Moore, M. H. (1997). *Creating public value: Strategic management in government*: Harvard university press.

Newman, J. (2014). Measuring policy success: case studies from Canada and Australia. *Australian Journal of Public Administration, 73*(2), 192-205.

Ojogiwa, O. T. (2021). The crux of strategic leadership for a transformed public sector management in Nigeria. *nternational Journal of Business*

Management Studies, 13(1), 83-96.

Parhizgari, A., & Gilbert, G. R. J. O. (2004). Measures of organizational effectiveness: private and public sector performance. *32*(3), 221-229.

Park, S., Lee, D. S., Son, J. J. R., & Governance. (2021). Regulatory reform in the era of new technological development: The role of organizational factors in the public sector. *15*(3), 894-908.

Parks, L., & Guay, R. P. (2009). Personality, values, and motivation. *47*(7), 675-684.

Pitlik, H. J. E. Ü., WIFO-Monatsberichte. (2017). Österreich 2025–Verwaltungsreform zwischen Effizienzstreben und Reformwiderständen. *90*(3), 205-217.

Pollitt, C. (2003). Joined-up Government: A Survey. *1*(1), 34-49. doi:10.1111/1478-9299.00004

Pollitt, C., & Bouckaert, G. (2017). Public management reform. In: Oxford University Press.

Putu S, N., Jan van Helden, G., & Tillema, S. (2007). Public sector performance measurement in developing countries: A literature review and research agenda. *Journal of Accounting & Organizational Change, 3*(3), 192-208.

Raaphorst, N., & Van de Walle, S. (2020). Trust, fairness, and signaling: Studying the interaction between officials and citizens. *The handbook of public sector communication*, 59-70.

Rainey, H. G. (2009). *Understanding and managing public organizations*: John Wiley & Sons.

Ravlin, E. C., & Meglino, B. M. J. J. o. A. p. (1987). Effect of values on perception and decision making: A study of alternative work values measures. *72*(4), 666.

Rhodes, R. A. (2016). Recovering the craft of public administration. *Public administration review, 76*(4), 638-647.

Rodriguez Müller, A. P., Martin Bosch, J., & Tangi, L. (2024). An overview of the expected public values arising from blockchain adoption in the European public sector. *International Journal of Public Sector Management*.

Sari, A. R. (2023). The impact of good governance on the quality of public management decision making. *Journal of Contemporary Administration and Management (ADMAN), 1*(2), 39-46.

Senyo, P. K., Effah, J., Osabutey, E. L. J. T. F., & Change, S. (2021). Digital platformisation as public sector transformation strategy: A case of Ghana's paperless port. *162*, 120387.

Simpson, H. J. J. o. E. S. (2009). Productivity in public services. *23*(2), 250-276.

Stainer, A., & Stainer, L. J. S. C. (1998). Strategic change in public services: a productivity and performance perspective. *7*(2).

Stoker, G. (2006). Public value management: A new narrative for networked governance? *The American Review of Public Administration, 36*(1), 41-57.

Suzuki, K., & Hur, H. J. P. M. R. (2020). Bureaucratic structures and organizational commitment: findings from a comparative study of 20 European countries. *22*(6), 877-907.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education, 2*, 53.

Ugyel, L., & O'Flynn, J. J. I. J. o. P. A. (2017). Measuring policy success: Evaluating public sector reform in Bhutan. *40*(2), 115-125.

Ukwandu, D. C., Ijere, E. S. J. A. s. P. S. D., & Review, P. (2020). Public sector reforms and national development: A situational analysis of post-independent Nigeria. *8*(1), 1-7.

Van Overwalle, F., & Siebler, F. J. S. C. (2013). A connectionist model of attitude formation and change. 345-423.

Wescott, C. G., & Jones, L. R. J. (2007). Managing for results and performance in Asia: Assessing reform initiatives in the public sector. *International Public Management Review, 8*(1), 56-102.

Zhang, E. J. A. H. (2021). Discourses on public sector accounting reforms in China: A brief history (1949–2019). *26*(2), 255-279.

Appendix:

Appendix 1

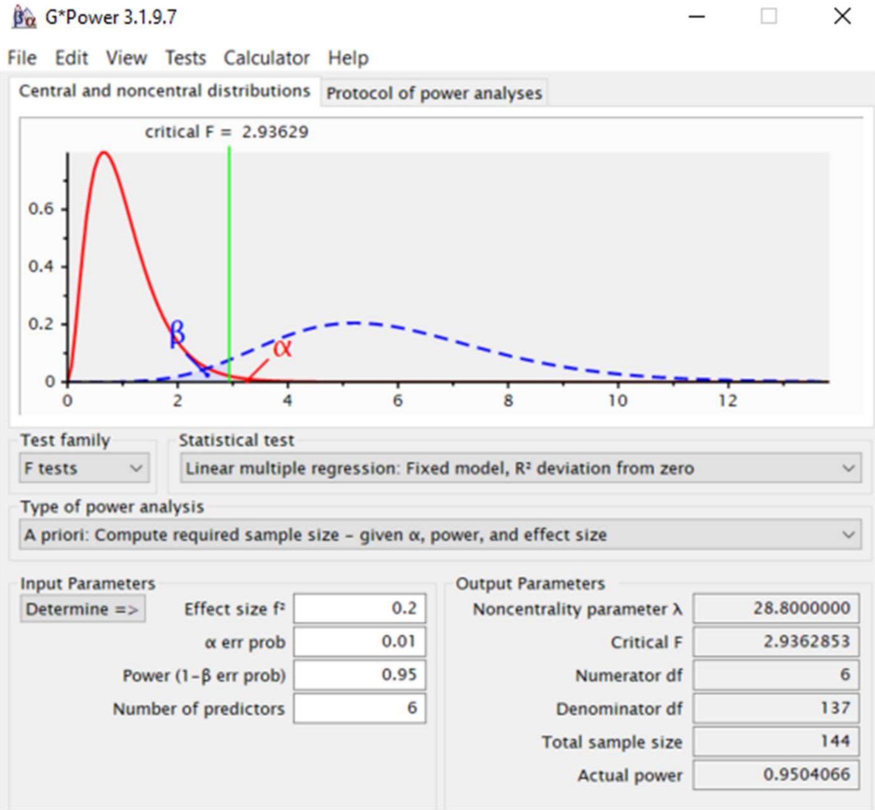


Table A1

Cronbach's alpha for the index "performance"

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation
Use of PI: Assess whether I reach my targets	24.59	97.119	.848
Use of PI: Monitor colleague performance	24.63	100.600	.806
Use of PI: Identify problems	24.27	98.149	.819
Use of PI: Learning and improvement	24.46	97.933	.859
Use of PI: Satisfy requirements of line manager	24.83	99.431	.792
Use of PI: Communicate org. contribution to users	25.32	104.567	.666
Use of PI: Engage with external stakeholders	25.54	105.315	.682
Use of PI: Manage image of org.	24.79	100.441	.768

Table A2

Cronbach's alpha for the paradigm indices

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation
TPA index			
Expertise	.0367	.035	.086
Rule of law	5.956	1.351	.086
NPM index			
Productivity	3.81	3.190	.495
Performance	3.91	2.569	.495

Table A3

Collinearity Statistics

Model		Tolerance	VIF
	Expertise	.931	1.074
	Rule of law (recoded)	.978	1,022
	Productivity	.752	1.329
	Performance	.757	1.321

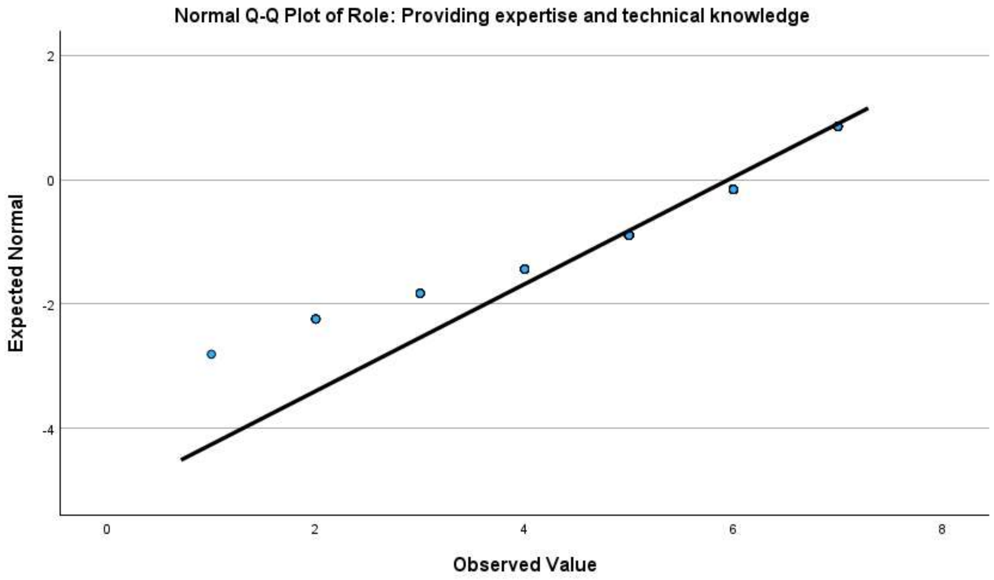
a. Dependent Variable: Reform success

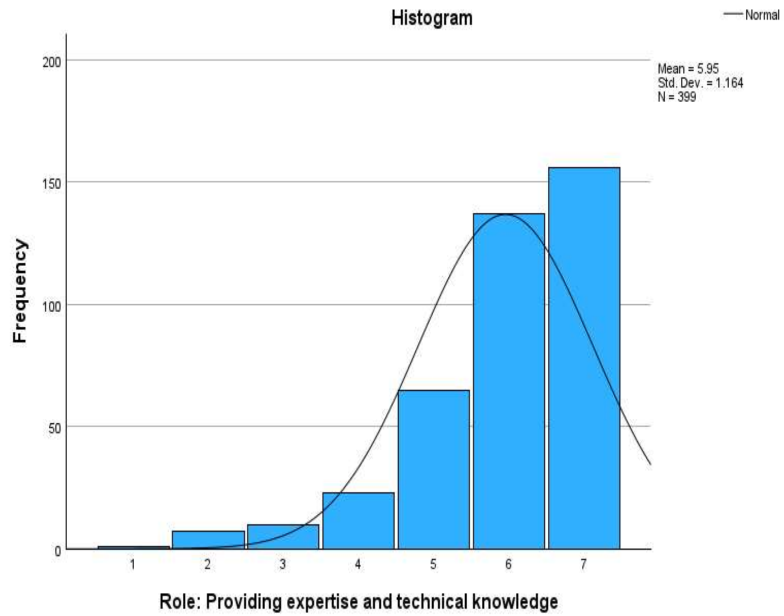
Table A4

Regression Diagnostics

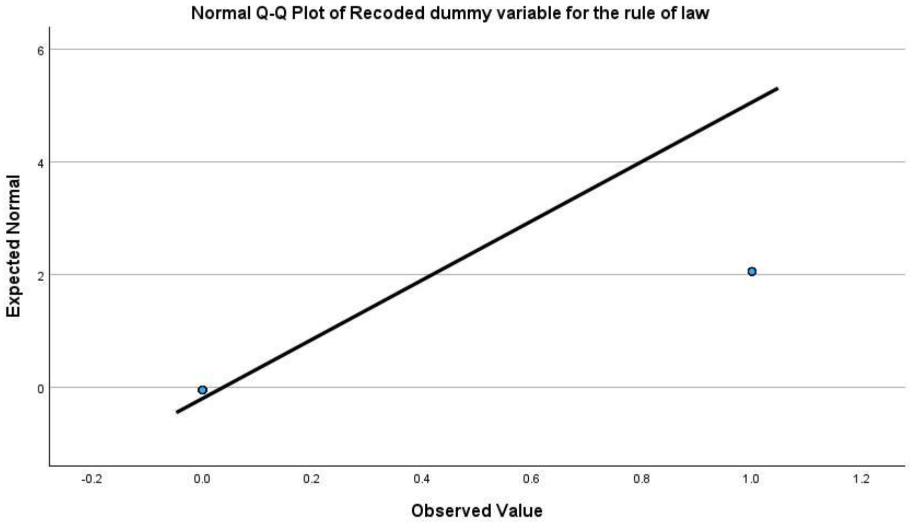
	Skewness		Kurtosis	
	Statistic	Std. Error	Statistic	Std. Error
Expertise	-1.365	.117	1.883	.233
Rule of law	4.982	.120	22.928	.239
Productivity	-.005	.116	-.975	.232
Performance	.029	.117	-1.052	.234
Reform success	-.236	.120	-.784	.240

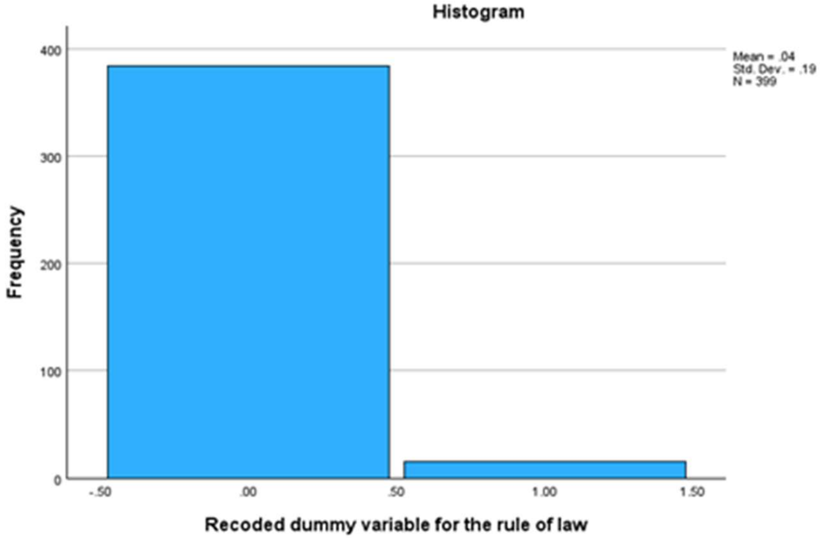
Appendix 2: Histogram and Q-Q Plot of "Expertise"



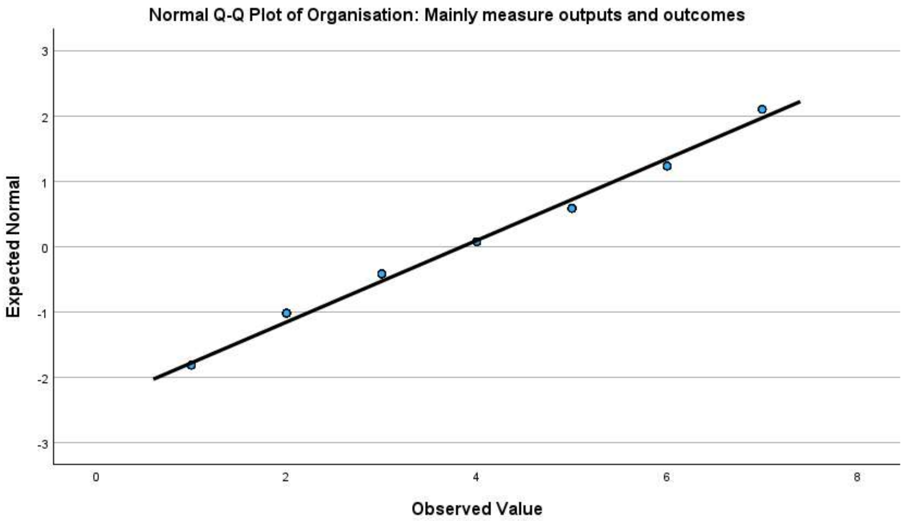


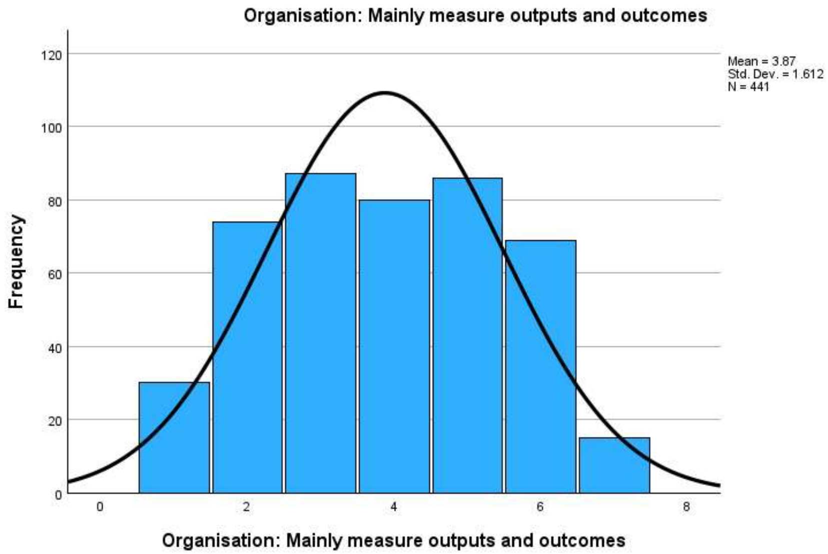
Appendix 3: Histogram and Q-Q Plot of "Rule of law"



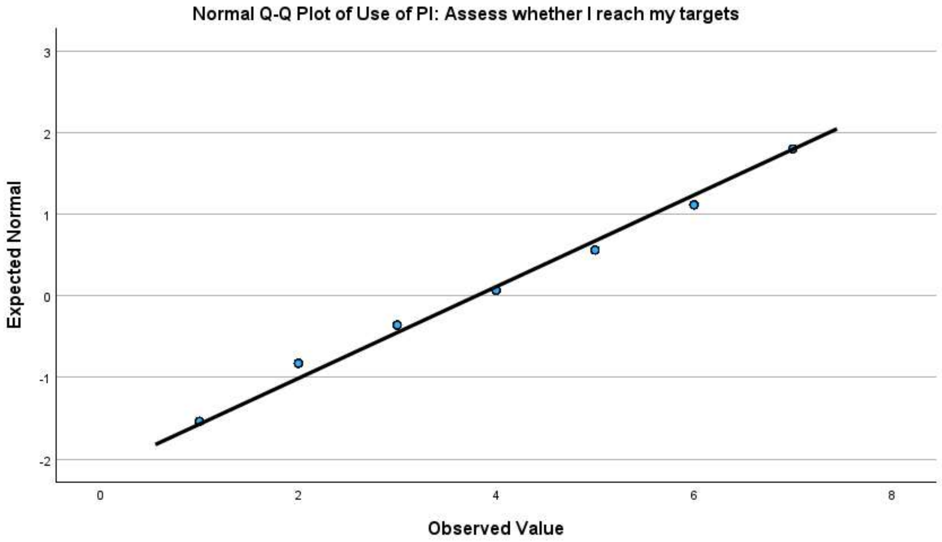


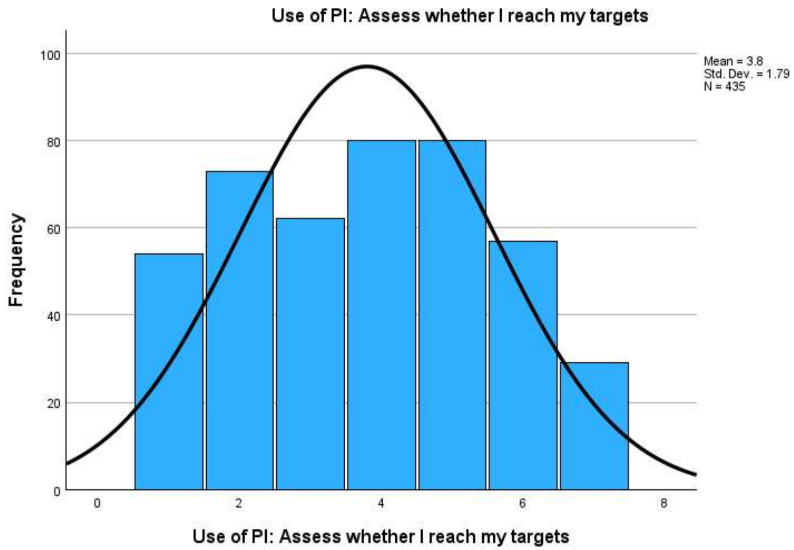
Appendix 4: Histogram and Q-Q Plot of "Productivity"





Appendix 5: Histogram and Q-Q Plot of "Performance"





Appendix 6: Histogram and Q-Q Plot of "Reform success"