The potential of implementing three beyond GDP frameworks: an evaluation from civil servants' perspective at the Municipality of Amsterdam

MSc MADE Thesis
21 November 2024
Amsterdam, the Netherlands

Author:
Mary-Ann van Kanten (1300660)
Mary-ann.vankanten@wur.nl

Wageningen University & Research supervisor:
Prof.dr.ir. Eveline van Leeuwen
Urban Economics
Eveline.vanleeuwen@wur.nl

TU Delft supervisor:
Dr. Karel van den Berghe
Management in the Built Environment
K.B.J.vandenberghe@tudelft.nl

Abstract

Beyond GDP approaches have trouble getting integrated into practice. This is however necessary in order to replace the 'economic-growth-GDP-compass' guiding government decision-making and to maximize their potential impact. The aim of the research was to gain insight into the user perspective on the implementation of beyond GDP frameworks in the city government context to inform the development of beyond GDP frameworks that have a higher potential for implementation to achieve deep system change. To this end, civil servants from the municipality of Amsterdam were asked to evaluate the Sustainable Development Goals, Doughnut Economics and Broad Prosperity frameworks for implementation. The frameworks were conceptualized as paradigm innovations that aim for change in the realm of intent of a system. An embedded mixed methods approach was taken, using a survey for the main quantitative approach and interviews for the secondary qualitative approach. Out of all the factors that influence innovation implementation, the focus was placed on the compatibility of the frameworks with the work of civil servants. Based on the research performed, it remains unknown which aspects of the frameworks were (not) considered in the minds of civil servants when they made the evaluations, but it is not unlikely that a degree of conceptual confusion underlies the outcomes.

Overall, the three frameworks are considered equally, slightly compatible. There are some differences between how subgroups of civil servants evaluate each of the frameworks depending on their work's relation to policy and the geographical scale that they work on. The extent to which civil servants' position in the organization explains the overall compatibility differs per framework. None of the frameworks excel in compatibility with any of five dimension of civil servants' work compared to the other dimensions. There are significant differences between the frameworks' compatibility within the dimensions of providing insight into considerations, internal communication and external communication. Compatibility with the five distinguished dimensions of civil servants' work explains significant but varying amounts of the variance in their overall compatibility. The explanatory power of compatibility with each of the individual dimensions remains unclear. Within the municipality of Amsterdam the majority of civil servants are either consistently in favor of the use of single or multiple frameworks throughout the organization, divided over similarly sized groups.

Key words

Beyond GDP frameworks, paradigm innovation, implementation, public organization

Table of Contents

1. Introduction

For decades, Gross Domestic Product (GDP) has been used as an indicator of economic progress and societal well-being, arguably even reducing societal well-being to this one indicator, guiding governments in their decision-making (Cavalletti & Corsi, 2016). Critiques on both the quality of the indicator for economic purposes and on its adequateness for representing well-being have been present since its development (Kuznets, 1934; Giovannini & Rondinella, 2018). Furthermore, GDP as an indicator represents frameworks of thought and action that conceptualize the economy as something linear for which growth and profit maximization is aspired, taking priority over all else (Raworth, 2017). Such frameworks of thought and action are arguably at the root of systems that produce complex issues like environmental degradation and social inequalities (Jackson, 2016; Raworth, 2017).

In recent years, both the academic world and society have increasingly recognized a need to shift the focus from economic growth measured in GDP and to 'go beyond GDP' by taking a multidimensional approach to well-being and societal progress. In other words, there is a sense of urgency for a paradigm shift away from the economic growth paradigm (Jansen et al., 2024). This has resulted in studies on alternative or complementary indicators for GDP (Stiglitz et al., 2009), as well as the development of multiple initiatives for the development of more holistic frameworks centered around sustainable and inclusive well-being (Jansen et al., 2024). Although governments now have a wealth of new frameworks and indicators to choose from, this does not automatically lead to the use of these new frameworks in practice. One reason for that is that classifications of indicators in academic literature have often not been relevant for practitioners to choose the right ones, which makes it difficult to integrate new indicators in practice (Bleys, 2012). The European Union BRAINPOoL project found a number of political, indicator, and process and structural barriers to the use of alternative indicators to GDP in policy making (Whitby et al., 2014).

Implementation science is a field of study that specifically studies the *uptake* of innovations into routine use, rather than their *impact* (Bauer & Kirchner, 2019). It has its roots in health sciences and emerged in response to the realization that most innovations that are tested to be effective in theory, do not make it into practice and therefore do not make their potential impact (Bauer & Kirchner, 2019; Implementation Science, n.d.). Effective implementation is a prerequisite for innovation effectiveness: only if an innovation is implemented successfully can the innovation potentially make the impact that an organization aimed to achieve from its use (Klein & Sorra, 1996; Real & Poole, 2004). While it remains arguable whether beyond GDP frameworks have been tested and proven to be effective, the focus on their uptake from theory to practice rather than the evaluation of their impact is in line with the research focus of implementation science. Therefore, the Implementation Science body of literature is considered appropriate to consult to further the knowledge on the implementation of beyond GDP frameworks in practice.

The work that lies before you now turns the focus towards the implementation of three beyond GDP frameworks at one organization: the Municipality of Amsterdam. The municipality of Amsterdam is an example of a governmental body that is looking for ways to go beyond GDP. In doing so, they are exploring the possibilities of three different frameworks at once: Sustainable Development Goals (SDGs), Doughnut Economics (DE) and Broad Prosperity (BP). The Sustainable Development Goals were developed by the United Nations and they were adopted in 2015 as the 2030 Agenda for Sustainable Development. The core message of the agenda is that "*Sustainable development recognizes that eradicating poverty in all its forms and dimensions, combating inequality within and among countries, preserving the planet, creating sustained, inclusive and sustainable economic growth and fostering social inclusion are linked to each other and are interdependent. (…) for today's generation and for future generations"* (United Nations, 2015, p. 5) The SDGs are presented a set of 17 goals, specified into 169 targets, that integrate and balance the three dimensions of sustainable development: the economic, social and environmental. Doughnut Economics was developed by Kate Raworth as *a radically new compass for guiding humanity this century* (Raworth, 2017). The Doughnut conceptualizes a social foundation and (global) ecological ceiling between which an ecologically safe and socially just space lies for humanity. She calls for a shift in thinking: *from "'good is forward- and- up' to 'good is in- balance'"* (Raworth, 2017, p. 53). The dimensions that make up the social foundation are inspired by the SDGs and the dimensions of the ecological ceiling are derived from Rockström et al. (2009). Broad Prosperity (BP) is a national and regional monitoring framework that has been developed by the Dutch National Bureau for Statistics (CBS). It is based on the outcomes of the Conference of European Statisticians Recommendations on Measuring Sustainable Development by the United Nations Economic Commission for Europe (UNECE/Eurostat/OECD, 2014). The CBS understands Broad Prosperity as balancing the quality of life here and now with the extent to which it deteriorates the quality of life of later generations and of people elsewhere in the world (Horlings & Smits, 2019). While the (thematic) operationalization into themes and indicators depends on what is relevant at the geographical scale in question, the meaning of Broad Prosperity remains the same.

Currently, the statistics department of the municipality of Amsterdam has incorporated the SDGs, The Doughnut and Broad Prosperity all at once in their two-yearly monitoring of how the city is doing on a multitude of topics (Gemeente Amsterdam, 2021). In other cases, one framework is chosen as leading: the DE framework is leading the city's circular economy strategy stemming from 2020 (Gemeente Amsterdam, 2020). Amsterdam was the first city to create policy based of the Doughnut (Amsterdam Doughnut Coalition, n.d.). In 2022 the city published its first Voluntary Local Review for the SDGs (Gemeente Amsterdam, 2022). At the same time, a project has started in the department of Digitization, Innovation and Information to research and develop a decision-making framework based on Broad Prosperity that can be used to assess the ex-ante impact of economic, ecological, spatial and socio-cultural policy. Two elements in this paragraph have not yet been studied: 1) whether it is (considered) problematic that departments do not all use the same frameworks and 2) whether it is problematic that single departments use multiple frameworks at the same time.

In all cases civil servants have to work with the frameworks that are chosen. Implementation research has found that committed and adequate use of an innovation, like the frameworks, is essential for their integration into the practice of an organization (Klein & Sorra, 1996). Without successful implementation, an innovation is unlikely to make its envisioned impact. In the case of the municipality of Amsterdam, no large scale systematic study has been performed to assess civil servants' perspective on the compatibility of the frameworks with their work related thoughts and actions. This then makes up the third and final point of interest for this thesis.

The following sections of this first chapter outline the problem statement, research aim and research questions guiding the study. An outline of the thesis report closes the chapter.

1.1 Problem statement

Beyond GDP approaches have trouble getting integrated into practice. This is however necessary in order to replace the 'economic-growth-GDP-compass' guiding government decision-making and to maximize their potential impact.

1.2 Research aim

The aim of this research is to gain insight into the user perspective on the implementation of Beyond GDP frameworks in the city government context to inform the development of Beyond GDP frameworks that have a higher potential for implementation to achieve deep system change.

1.3 Research questions

The main research question is:

How do civil servants of the municipality of Amsterdam evaluate three beyond GDP frameworks (The Sustainable Development Goals, Doughnut Economics and Broad Prosperity) for implementation?

The sub-questions are:

1. How do civil servants evaluate the frameworks' overall compatibility?
2. To what extent can civil servants' evaluations of the overall compatibility of the three frameworks be explained by their position in the organization?
3. How do civil servants evaluate the frameworks' compatibility with five dimensions of their work?
4. What are the relationships between the evaluations of five dimensions of compatibility and the overall compatibility score?
5. What are civil servants' attitudes towards the use of single or multiple frameworks throughout the organization?

1.4 Report structure

The following chapter (2) describes the theoretical framework that underlies the research endeavors. It includes a conceptualization of beyond GDP frameworks, an elaboration on paradigm change, innovation types, and innovation implementation. Next is the methodology chapter (3), explaining the steps taken and decisions made regarding data collection and analysis. The results (4) are presented structured by sub-question. Then the results are discussed (5) with reflections on the methods used, the literature and the bigger research context. Finally, the report ends with the conclusions, policy recommendations and suggestions for future research.

2. Theoretical Framework

2.1 An Understanding of Beyond GDP Frameworks

At the center of this research are three 'beyond GDP frameworks'. The term 'framework' can be used in many different ways. Think about the different kinds of frameworks that are used in academic research: theoretical frameworks, conceptual frameworks and analytical frameworks, to name a few from the top of my head. They all give some kind of structure to something, yet they are all conceptually distinct. So what does it mean to speak of 'beyond GDP frameworks'?

Boarini & D'Ercole (2013) have titled the second section of their paper "*The OECD framework for measuring well-being*". In this section they present the conceptual framework that the OECD uses for defining and measuring well-being. In a footnote, they further specify how they understand 'framework' in this case: "*a graphical representation of the dimensions and items that should be taken into account for any comprehensive assessment of individual and societal well-being*" (Boarini & D'Ercole, 2013). They consider the framework an attempt to operationalize the capabilities approach to well-being.

Giovannini & Rondinella (2018) provide an evaluation of the development of some of the theoretical and conceptual approaches of going beyond GDP and methodological challenges since the 1930's. They state that there has been a demand for "*an* [analytical] *framework to assess social and economic conditions of citizens and a* [theoretical] *approach to measure them*" (Giovannini & Rondinella, 2018, p. 6). From their evaluation, it is possible to extract that beyond GDP frameworks are concepts, such as development, welfare and well-being, that have been conceptualized and operationalized to a certain extent and in a certain way. Frameworks are developed by different entities (e.g. research institutes, (inter)governmental organizations, economic organizations), informed by theoretical approaches and with certain use purposes in mind.

Rather than concepts, indicators are the starting point for Bleys (2012). He argues that (new) alternative indicators should be classified in a way that helps policy makers decide which indicators are most suitable for them to use. The classification he proposes starts by tracing back the concepts that indicators are attempting to quantify. From this slightly different perspective, beyond GDP frameworks can be considered as organizations of alternative indicators. So rather than the operationalization of concepts, it is a classification of indicators.

Combining the three perspectives above, results in the understanding of 'what' beyond GDP frameworks are that will be used in this thesis. A beyond GDP framework is a conceptual framework and its operationalization. The logic behind the framework may either be in the direction from concept towards indicator or the other way around, starting with indicators that are classified using increasingly higher levels of conceptualization.

2.2 Paradigm Innovation

This study considers the Sustainable Development Goals, Doughnut Economics and Broad Prosperity frameworks as new ideas, new ways of thinking about and measuring 'what really matters beyond economic growth measured in GDP' that may be used by a public organization. Considering the newness of the frameworks for the organization and the implementation gap that was identified, let us turn towards innovation (implementation) literature for guidance.

At the organizational level, innovation can be defined as "*the development (generation) and/or use (adoption) of new products, services, technologies, systems or practices*" (Damanpour, 2006, p. 216). To this definition, Baregheh et al. (2009) add that this is a multi-stage process that is undertaken with the aim of achieving some specific kind of progress. The new thing itself can be referred to as *an* innovation. For example, Wisdom et al. (2014) examine the adoption *of* innovations in the process of innovation implementation. Similarly, when Damanpour & Schneider (2006) discuss the implementation phase of process of innovation adoption they speak of "*modifying* [the] *innovation, preparing the organization for* [its] *use, (…)*" (p. 217).

It has been argued that it is useful to make a distinction between different types of innovation, and public and private organizations to better understand the drivers and barriers to their use in practice. The type of an innovation can be determined in a multitude of ways, but there are two main approaches. The first is to base innovation types on the outcomes of an innovation process as can be seen in the approaches of Geels et al. (2004) and Baregheh et al. (2016). Since the implementation process of the three beyond GDP frameworks at the center of this research is not yet finalized, a typology that is based on the initial focus of change is used. Buchheim et al. (2019) distinguish between four types of innovation in the public sector based on their initial focus of change: product or service, process, governance and paradigm innovations. Product or service and process innovations have received most attention in research and are self-explanatory in their focus. The focus of change of governance innovations is the broad social conditions or institutional framework in which public sector organizations operate (Moore & Hartley, 2008). Paradigm innovations focus on changing the mental frameworks or perspectives through which an organization approaches its issues (Francis and Bessant 2005). De Vries et al. (2016) use the category ´conceptual innovation' instead of 'paradigm', but it is not far removed in terms of its focus: "*introduction of new concepts, frames of reference or new paradigms that help to reframe the nature of specific problems as well as their possible solutions*" (p. 153). Since the focus of change of beyond GDP frameworks is on a shift away from the economic-growth-GDP paradigm by introducing new conceptualizations and operationalizations of well-being and social progress, they can be understood as paradigm innovations. Paradigm innovations are amongst those that have received the least attention in innovation literature (Buchheim et al., 2019).

While the literature on innovation may lack attention for paradigm innovation, the concept of paradigm shifts is more prominent in the field of systems thinking. Donella Meadows (1999) introduced the concept of leverage points: places within a complex system (e.g. corporation, economy) where a small shift in one thing can change the system as a whole. The leverage points are at different levels, ranging from shallow to deep. Interventions at shallow leverage points are relatively easy to make, but they have limited potential of leading to transformative change. On the contrary, interventions at deeper levels have a high potential of leading to transformative change, but are difficult to implement. Abson et al. (2016) aggregated the leverage points into four types of system characteristics that can be the target of intervention: parameters, feedback loops, design and intent. At the level of intent, the objective is to change the *norms, values and goals embodied within the system of interest and the underpinning paradigms out of which they arise* (Abson et al., 2016, p. 32). Deeper level system characteristics constrain which interventions may be taken at more shallow leverage points. From least to most deep, the three leverage points in the realm of intent are 1) the goals of the system, 2) the mindset or paradigm out of which the system arises and 3) the power to transcend paradigms.

2.3 Innovation implementation

The process of the uptake of innovations at organizations in practice can be divided into stages. Klein & Sorra (1996) make the distinction between source-based stage models and user-based stage models for the innovation process. They differ in perspective: the former type is based on the perspective of the innovation developer, while the latter type is based on the perspective of (intended) users of the innovation. The latter type would be relevant to this study as it can accommodate the perspective of civil servants, who would be the users of one or more of the three frameworks. Still within this type of stage model, authors use terminology slightly differently. From the user perspective, Klein & Sorra (1996) divide the innovation process in the stages of awareness, selection, adoption, implementation and routinization. They consider the innovation adoption as a necessarily preceding stage for the implementation stage. Aarons et al. (2010) identify the stages of exploration, adoption/preparation, implementation and sustainment as parts of the innovation implementation process. Damanpour & Schneider (2006) and Damanpour & Schneider (2008) divide the innovation adoption process into the stages of initiation, adoption decision and implementation. While the terms attached to the full process and some stages differ, a common thread is that the implementation stage follows the (pre-)adoption stage(s). The adoption stage ends with the decision to fully or partially implement the innovation (Wisdom et al., 2013). Although the full process of the uptake of innovations is far from linear in reality and stages can overlap, thinking along the lines of a rough time scale in which certain actions and stages precede or follow one another is still useful for analytical purposes (Real & Poole, 2005).

Further specifying the *uptake* of an innovation at an organization, the implementation stage at an organization can be defined as "(…) *the process of gaining targeted employees' appropriate and committed use of an innovation*" (Klein & Sorra, 1996, p. 1055). They define targeted employees as those who will use the innovation directly and those who are expected to supports its use. In the beyond GDP implementation context, research has mostly focused on the appropriate use of indicators. Hayden (2021) finds three uses of beyond GDP indicators linked to the policy process: political, conceptual and instrumental use. Indicators can be used by political actors to justify policy positions or to convince others of their viewpoints. In that case, the use of beyond GDP indicators does not necessarily extend to the realm of actively altering the policies that are being produced. Indicators can be used conceptually to create new understandings, narratives or frameworks of thought about what the focus should be on if not on economic growth measured in GDP. When indicators are directly linked to policy decisions, their use is considered instrumental. In the conceptualization of beyond GDP frameworks presented in the first section of this chapter, indicators are considered an element of the whole that already embodies a conceptualization of what the focus should be on if not economic growth measured in GDP.

The implementation stage may include actions towards modifying the innovation, preparing the organization for its use and trial use (Damanpour & Schneider, 2006). In the cases of the SDGs and Doughnut Economics, such activities have been considered necessary in the form of localization (Ciambra et al., 2023) or downscaling (Turner & Wills, 2022) to make the frameworks fit the activities performed at lower geographical scales. Hayden (2021) shares a number of possible processes and product through which implementation could take place for beyond GDP indicators. When the implementation endeavors result in consistent use of an innovation by targeted individuals in the organization, implementation is considered effective (Klein & Sorra, 1996) or successful (Kim & Chung, 2017).

Multiple studies have (re-)structured the factors that were found to be determinants for implementation through different case studies. It is found that the exact factors that influence innovation implementation are very specific to an organization, even compared to organizations within the same industry (Klein & Sorra, 1996; Aarons et al., 2010). Yet, types of factors can be identified. There are multiple ways of classifying the factors that affect implementation: contextual factors vs. innovation characteristics (Damanpour & Schneider, 2008), contextual factors within an organization vs. outside of the organization (Aarons et al., 2010), and organizational level vs. individual level factors (Singh et al., 2020). These categories may overlap. Kim & Chung (2017) divide contextual factors in the categories of social, organizational, and individual factors. For public sector innovations, De Vries et al. (2016) define four main categories on four levels:

- environmental level: external context (e.g. political mandates)
- organizational level: aspects that include the structural and cultural features of an organization (e.g. organizational slack resources)
- innovation level: intrinsic attributes of an innovation (as perceived by prospective adopters) (e.g. compatibility of the innovation)
- individual/employee level: characteristics of individuals who innovate(e.g. empowerment).

The scope of this study is set to innovation level characteristics of beyond GDP frameworks. Rogers (2003) distinguishes six innovation characteristics that are important to consider for implementation: relative advantage, compatibility, complexity, trialability, observability and re-invention potential. Compatibility as perceived by prospective users is particularly interesting in the case of paradigm innovation. Compatibility is defined as "*degree to which an innovation is perceived as being consistent with the existing values, past experiences, and needs of potential adopters. An idea that is not compatible with the prevalent values and norms of a social system will not be adopted as rapidly as an innovation that is compatible. The adoption of an incompatible innovation often requires the prior adoption of a new value system"* (Rogers, 2003, p. 15). Since the focus of paradigm innovation is to fundamentally change the mindset out of which the organizational system arises, it would logically follow that paradigm innovations are not considered very compatible.

Korteland & Bekkers (2008) find that innovation in the context of the public sector is characterized by not only using a logic of consequence, like in the private sector, but also a logic of appropriateness. According to them this means that the implementation of an innovation is not only influenced by whether an innovation is more effective and efficient than a certain reference point, but that it also matters how the innovation relates to the political environment in which the organization operates and whether it is desirable or necessary to 'match' with other organizations, referred to as isomorphism. While they categorize compatibility as stemming from a logic of consequence, this factor could arguably also capture innovation compatibility with the political environment and with other organizations.

3. Methodology

This chapter outlines the methodological choices that were made and provides a recipe for repetition. First, a description of the context of this research is provided, followed by the research design and epistemological and ontological approaches. Next, the data collection and analysis methods are outlined. Finally, the implications of the choices for this research's reliability and validity are discussed.

3.1 Research context

3.1.1 The case: Municipality of Amsterdam

At the center of this study is the organization of over 19,000 civil servants of the municipality of Amsterdam, the Netherlands. At this city government, the organization of civil servants operates to achieve the ambitions of the city that are set out in the political coalition agreement. Such an agreement is made every four years after local elections. The latest coalition agreement was set for 2022-2026. It is structured in four chapters capturing the social, sustainability, economic, spatial and governance dimensions of policy. This edition states that the Broad Prosperity is the starting point for the development of a future proof economic structure for the city (Gemeente Amsterdam, 2022, p. 47). In addition, they state that a new vision on the local economy will be created with broad prosperity as the goal. The civil servants that translate these ambitions into policies, programs and projects are divided among five clusters, which are further divided thematically into departments and smaller teams (**figure 1**).

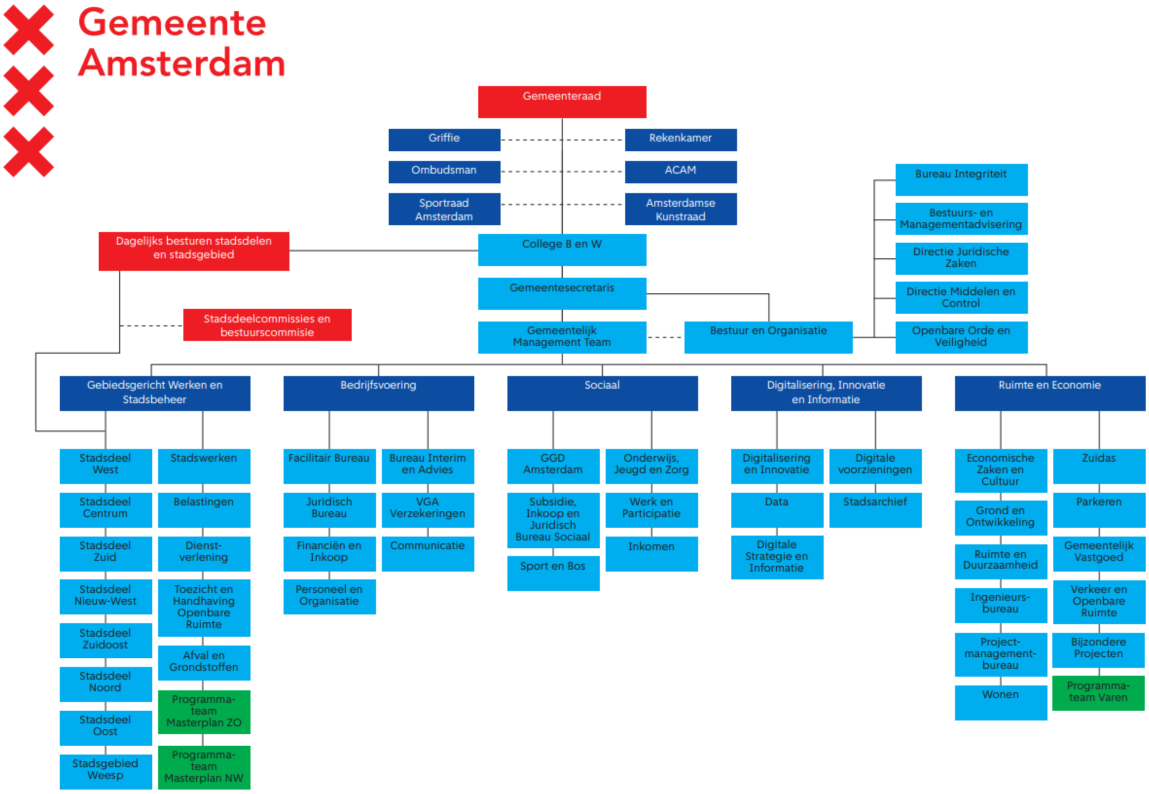


***Figure 1**: organizational structure of the municipality of Amsterdam. Red = political entities, dark blue = clusters, light blue = departments, green = special programs that cross the borders of departments and clusters.*

3.1.2 AMS Institute Ideal(s) City project

The research for this thesis has occurred partly within the context of the Ideal(s) City project of the AMS Institute. The AMS Institute is a research institute established by Wageningen University and Research, TU Delft and MIT. It collaborates closely with the Municipality of Amsterdam to provide solutions to urban challenges.

The purpose of the Ideal(s) City project is to develop a framework for the municipality of Amsterdam (in collaboration) to monitor the ideals of the city by using the already existing monitoring frameworks as building blocks. Ideals, goals, frameworks and indicators should be aligned with each other in this. This thesis functions as a pilot version of a larger scale study of the evaluation of beyond GDP frameworks by civil servants. The established contacts of the project team with civil servants were called upon. In the following paragraphs, it is mentioned when and how they were used, as well as other instances in which the project has played a role.

3.2 Research Strategy and Design

To answer the research question(s), an embedded mixed methods research design has been deployed in which the quantitative takes the dominant position. The reason for taking a quantitative approach is that the goal has been to make claims that can be generalized to the organization of 19,000 civil servants – descriptive claims that indicate where there are differences between opinions throughout the organization and relationships between pre-determined variables. To this end, opinions and attitudes have been quantified. A secondary qualitative approach is taken to capture potential rationales behind the quantification of opinions and attitudes in an exploratory fashion. The researcher assumes that the concepts under study are open to interpretation of the research participants. The qualitative approach then captures some first rationales that can help towards the interpretation of the outcomes following from the quantitative approach.

3.3 Data collection

3.3.1 Survey

The main method of data collection is a survey that was used was developed as part of the Ideal(s) City project and therefore contains a few questions that were relevant to other parts of the project, but not this exact study. The following paragraphs discuss the content of the survey that is relevant to this study, how the survey has been developed and its dissemination. Appendix A can be consulted for the full list of questions included in the survey.

Variables

Table 1 provides an overview of the variables that are considered in this study. Sub-questions 1 to 4 regard the compatibility of the three frameworks with respondents' work. Compatibility is understood as it is defined by Rogers (2003): "*the degree to which an innovation is perceived as being consistent with the existing values, past experiences, and needs of potential adopters*" (Rogers, 2003, p. 15). Respondents were asked to evaluate each framework's compatibility with their work in two ways. First, they were asked to rate each framework's suitability for guiding their work in general on a 1 to 10 point scale. A score of 1 indicates that the framework is not at all suitable, a score of 10 indicates that the framework is very suitable. Next, compatibility with needs has been further operationalized as the degree of suitability of using each of the frameworks for five dimensions of the work that civil servants have to perform. The first dimension, goal formulation, stems from the idea that civil servants work towards the achievement of a range of goals. If beyond GDP frameworks should replace the economic-growth-GDP framework that goals now follow from, they need to also be compatible with formulating goals. The second dimension, providing insight into considerations, is derived from the fact that Broad Prosperity is specifically being developed for this purpose: this indicates that there is a need here for beyond GDP frameworks to fulfill. The third dimension, evaluating results, follows from the idea that dimensions and indicators should be able to capture the criteria that determine the extent to which goals are achieved. The fourth and fifth dimensions of internal and external communication follow from the fact that civil servants do not work in isolation and that they have to be able to justify their decisions using appropriate rationales and indicators, both within the organization and beyond. In a matrix format, respondents were asked to rate each of the frameworks as 1) not suitable, 2) somewhat suitable or 3) very suitable for each of the five dimensions.

Considering the size of the organization and the plethora of tasks that are performed, three characteristics of the position of civil servants in the organization were tested for their influence on the evaluation of framework compatibility on a 1 to 10 point scale. These characteristics are the organizational cluster that they work for, the relationship of their work with policy and the geographical focus of their work. The organization of civil servants at the municipality of Amsterdam is divided in five thematical clusters: Gebiedsgericht Werken and Stadsbeheer, Bedrijfsvoering, Sociaal, Digitalisering, Innovatie and Informatie, Ruimte and Economie.

Considering that the tasks of the organization of civil servants revolve around policy, the position that respondents' work has in the policy process has been derived from the bottom half of the policy cycle of Seaford (2013). Respondents could indicate whether their work is concerned with policy making, implementation and/or evaluation. Respondents were asked one question in which could select all options that applied. The variable was then recoded so that 0 = not involved in policy making, implementation and/or evaluation, 1 = involved in policy making, implementation and/or evaluation.

Table 1
Variable overview

Variable	Dimension	Measurement
Compatibility with work related needs: the suitability of a framework for guiding tasks that have to be performed by civil servants.	Overall	1 to 10 point scale
	For goal formulation	3-point Likert Scale
	For providing transparency in considerations	3-point Likert Scale
	For evaluation of results	3-point Likert Scale
	For internal communication	3-point Likert Scale
	For external communication	3-point Likert Scale
Organizational cluster	Gebiedsgericht Werken and Stadsbeheer	
	Bedrijfsvoering	
	Sociaal	
	Digitalisering, Innovatie and Informatie	
	Ruimte & Economie	
Work's relation to policy	Policy making	0 = not related, 1 = related
	Policy implementation	0 = not related, 1 = related
	Policy evaluation	0 = not related, 1 = related
Geographical focus	< city	0 = not focused on, 1 = focused on
	City	0 = not focused on, 1 = focused on
	> city, within the country	0 = not focused on, 1 = focused on
	International	0 = not focused on, 1 = focused on
Attitudes towards the use of single or multiple frameworks	*It is best to use one framework for the entire process of policy development, implementation, evaluation and communicating about it.*	Agreement on a 3-point Likert Scale
	Which framework is suitable, depends on the policy theme.	Agreement on a 3-point Likert Scale
	It is problematic that different departments use different frameworks.	Agreement on a 3-point Likert Scale

The operations of the municipality as an organization are concerned with different geographical scales, both within their own territory and beyond. The scales 'smaller than the city scale', 'city', 'beyond the city, but within the country' and 'international' were distinguished. Similar to civil servants' relation to policy, respondents were asked in a single question to select all geographical scales that apply to their work. The variable was then coded do that 0 = not focused on a certain scale, and 1 = work is focused on a certain scale.

The fifth and final sub-question addresses civil servants' attitudes towards the use of single or multiple frameworks throughout the organization. They were asked to indicate their (dis)agreement to three statements on a 3-point Likert Scale with the option of checking an 'I don't know' box. The statements were:

1. *It is best to use one framework for the entire process of policy development, implementation, evaluation and communicating about it.*
2. *Which framework is suitable, depends on the policy theme*.
3. *It is problematic that different departments use different frameworks*

Development process

The survey has been developed using Qualtrics survey software and has been written in Dutch, because this is the main language used at the municipality of Amsterdam. The survey was informed by an introductory conversation with one civil servant who is involved in the development of the Broad Prosperity framework in the municipality of Amsterdam and three rounds of feedback: 1) with the same civil servant from the introductory conversation, 2) with one of their colleagues who has more experience with surveying within the organization of the municipality and 3) a test round of the survey with four other civil servants. These feedback rounds were used to improve the understandability of the questions, to take into consideration any political sensitivities surrounding the subject that may influence people's willingness to complete it, and to test for its duration.

Dissemination

The aim was to distribute the survey as broadly as possible throughout the organization. It was recommended to send the survey request by e-mail via familiar contacts and to link the research to the AMS Institute project 'The Ideal(s) City' to increase the chances of a higher response rate. Therefore, the survey has been disseminated by using the method of snowballing. Contacts from the first two civil servants who were involved in the development of the survey and contacts from The Ideal(s) City project were the starting points. The majority of the first line of contacts were part of Ruimte & Economie, Gebiedsgericht Werken and Innovation. In addition, three posts were placed on LinkedIn at the end of May and the beginning of June by members of the 'Ideal(s) City' project team.

3.3.2 Interviews

Table 2
Overview of interviewees

Interviewee	Framework	Contact	Setting	Duration
1	Broad Prosperity & Doughnut Economics	Via Ideal(s) City project	Online	1 hour
2	Frameworks in general, Sustainable Development Goals, Broad Prosperity, Doughnut Economics	Approached the researcher after completing the survey	In-person	1.5 hours
3	Broad Prosperity & Sustainable Development Goals	Via interviewee 1	In-person	35 minutes
4	Doughnut Economics & frameworks in general	Via Ideal(s) City project	Online, together with two Ideal(s) City project members	1 hour

In total, four interviews were conducted. Interviewees were sought with the intent of interviewing one person involved in the implementation of each of the frameworks. They are assumed to be the most knowledgeable on the topic within the context of the municipality of Amsterdam. As such, they 1) can provide insight into their own rationales and 2) they are assumed to possess knowledge about the rationales of other civil servants throughout the organization through their experience with the implementation of the frameworks. The interviewees will remain anonymous throughout this report and have therefore been given numbers to distinguish them. Table 2 outlines the frameworks that were discussed in each interview, how interviewees were contacted and the interview setting.

The interview item list contained three topics. The first topic was the appropriate use and targeted users that are envisioned for the frameworks. This forms a counterpart to the survey, which asks targeted users about their judgement of the frameworks' compatibility with their work. With the aim of this research being to contribute knowledge to the design of frameworks in mind, the second topic asks specifically about framework properties that form barriers or opportunities for implementation. Considering the knowledge gap of the implementation of multiple innovations at the same time at the same organization, the third topic concerned competition and complementarity between the three frameworks.

3.4 Analysis

3.4.1 Sample under analysis

Due to the potential that respondents were unfamiliar with one or more of the frameworks and the large number of individual questions contained in the matrix, respondents could leave boxes empty if they wished to do so. As a consequence of this design choice, not every respondent in the dataset answered each question that is necessary to answer the research questions. This means that 1) the sub-sample sizes per variable can differ, 2) the sub-samples can differ in composition/represent different populations and 3) that the same respondents can be part of multiple sub-samples. This can become problematic when making comparisons between frameworks: analysis outcomes can be the result of sample differences rather than actual differences caused by the variables of interest. Moreover, statistical analyses are generally equipped to either work with samples comprised completely of paired observations or completely independent observations (Derrick et al., 2017). In the case of this study, the decision was made to perform analyses on two sub-samples. The first four sub-questions are focused on the evaluation of the compatibility of the three beyond GDP frameworks. The analyses to answer these questions have been performed on the sub-sample of respondents who scored all three of the frameworks on overall compatibility and all five dimensions of compatibility. The fifth sub-question is answered using the group of respondents that responded to all three of the statements about the use of single or multiple frameworks.

3.4.2 Methods

The analysis of the quantitative data took place in SPSS. All throughout, a significance level of $p < 0.10$ was accepted, unless stated otherwise. The sections below outline the analysis steps taken to answer the sub-questions.

1. *How do civil servants evaluate the frameworks' overall compatibility?*

The analysis to answer the first sub-question looked at the evaluation of the overall compatibility scores of the frameworks on a 1 to 10 point scale. First, the descriptive statistics of the total sample are calculated. Then, a Friedman test was performed to test for differences between the overall compatibility scores of the three frameworks. Then, Mann-Whitney U tests were performed to determine whether there are differences between how groups within the organization scored the frameworks' overall compatibility with their work. The organizational cluster that civil servants work at, their relation to policy and geographical focus formed the basis upon which groups were defined. When there differences between groups were found, the means were compared.

2. *To what extent can civil servants' evaluations of the overall compatibility scores on a 1 to 10 point scale be explained by their position in the organization?*

To answer the second sub-question, a linear regression analysis was performed to test the extent to which civil servants' position in the organization explains the overall compatibility score they give to each of the frameworks. The independent variable is the overall compatibility score on a 1 to 10 point scale. The independent variables are the binary coded variables for organizational cluster, the relation of civil servants' work to policy making, implementation and evaluation, and the geographical focus of their work being on the scale smaller than the city, the city scale, beyond the city within the Netherlands and the international scale. After calculating Spearman's rho correlation coefficients, the variables were added block-wise in a multiple linear regression for each of the frameworks to assess the explanatory power of each of the variables. Cluster was entered for the first models, work's relation to policy is entered to create the second and the geographical focus was added third for the final models. The models were checked for collinearity. VIF of maximum 10 was considered acceptable. This resulted in the calculation of the following three models for each of the three frameworks:

Model 1: *Framework*$_n$ *compatibility score on a 1 to 10 point scale = βCluster Ruimte&Economie*

Model 2: *Framework*$_n$ *compatibility score on a 1 to 10 point scale = βCluster Ruimte&Economie + βpolicy making + βpolicy implementation + βpolicy evaluation*

Model 3: *Framework*$_n$ *compatibility score on a 1 to 10 point scale = βCluster Ruimte&Economie + βpolicy making + βpolicy implementation + βpolicy evaluation + β<city + βcity + β>city within the country + βinternational*

3. How do civil servants evaluate the frameworks' compatibility with five dimensions of their work?

The third part of the analysis starts with comparing the scores that respondents assigned to framework compatibility with five dimensions of their work on a 3-point Likert Scale. First the evaluations are described using the descriptive statistics. Then, a total of eight Friedman tests were performed to test for differences between the evaluations of the frameworks per dimension (5 tests), as well as differences between the scores for the dimensions per framework (3 tests). When the results indicated differences, post hoc Wilcoxon tests were performed to find out which evaluations differ statistically. A Bonferroni adjustment was made to correct for the fact that multiple Wilcoxon tests were performed, which led to an accepted significance level of $p < 0.10$/the number of comparisons being made.

4. To what extent does the framework compatibility with five dimensions of civil servants' work explain the overall compatibility scores of the frameworks?

For the fourth step of the analysis, multiple regression analyses were run to assess the extent to which the scores that respondents assigned to framework compatibility with five dimensions of their work on a 3-point Likert Scale explain the overall compatibility scores on a 1 to 10 point scale. The independent variables were all added at once. Again, VIF of maximum 10 was considered acceptable. For each of the three frameworks the following model was calculated:

Model: *Framework*$_n$ *compatibility score on a 1 to 10 point scale = βgoal formulation score + βproviding insights into considerations score + βevaluating results score + βinternal communication score + βexternal communication score*

5. What are civil servants' attitudes towards the use of single or multiple frameworks?

The fifth and final part of the analysis regards the attitudes of civil servants towards the use of single or multiple frameworks throughout the organization. After assessing the descriptives and frequencies of responses, a correlation analysis using Spearman's rho was performed to assess whether civil servants' attitudes towards the three statements are more or less aligned or contradicting. When significant correlations were found, crosstabulation was used to assess the sizes of the groups with aligned or contradicting attitudes towards the statements.

The interviews have been used to interpret the results in the first section of the discussion. The interview transcripts have not been coded to this end. Instead, the researcher's judgement was used to determine whether interview data was related to potential reasoning behind the quantified opinions and attitudes. Transcripts can be made available on request.

4. Results

This chapter presents the results of the data analysis as described in the previous chapter. It is structured along the sub-research questions set out at the beginning of this report. Once more, they are:

1. How do civil servants evaluate the frameworks' overall compatibility?
2. To what extent can civil servants' evaluations of the overall compatibility of the three frameworks be explained by their position in the organization?
3. How do civil servants evaluate the frameworks' compatibility with five dimensions of their work?
4. What are the relationships between the evaluations of compatibility with the five dimensions and the overall compatibility score?
5. What are civil servants' attitudes towards the use of single or multiple frameworks throughout the organization?

4.1 Overall compatibility scores of the SDGs, Doughnut Economics and Broad Prosperity

The descriptive statistics of the overall compatibility scores of the frameworks on a 1 to 10 point scale are presented in table 3 and the distributions are visualized in figure 1. Broad Prosperity has received the highest average compatibility score of 6.65 out of 10 (SD = 1.924, var = 3.703). The Doughnut comes in second place with a score of 6.06 (SD = 2.205, var = 4.862). The SDGs score lowest with an average overall compatibility score of 5.94 (SD = 1.982, var = 3.929). The outcome of a Friedman test is that there is no statistical difference between the overall compatibility scores of the three frameworks, $\chi^2(2) = 3.651$, $p = 0.161$.

Table 3
Descriptive statistics for the scoring of the Sustainable Development Goals, Doughnut Economics and Broad Prosperity frameworks on a 1 to 10 point scale.

Framework	Mean	SD	Variance	Minimum	Maximum
SDGs	5.94	1.982	3.929	2	9
DE	6.06	2.205	4.862	2	10
BP	6.65	1.924	3.703	3	10

Note: N=31 for all three frameworks.

Figure 1: Frequencies of Overall Compatibility Scores of SDGs, Doughnut, and Broad Prosperity on a 1 to 10 Point Scale

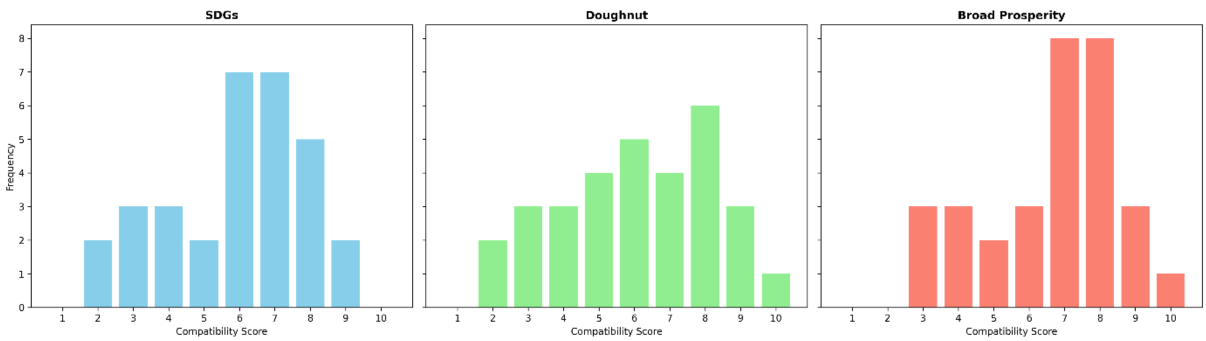


Table 4 presents the statistically significant Mann-Whitney U test results for differences between how groups throughout the organization scored the compatibility of the three frameworks on a 1 to 10 point scale and the associated means. Please consult Appendix B for the full set of outcomes from the Mann-Whitey U tests. The results indicate that civil servants whose work is related to policy making give the SDGs a lower average score (5.25) than those whose work is not (6.37). Similarly, those whose work is focused on the city scale give the SDGs a lower average compatibility score (5.56) than those whose work is not focused on this geographical scale (7.50). On the other hand, respondents who work is focused on the international scale give the SDGs an average overall compatibility score of 7.14 compared to the lower score of 5.58 of those not focused on the international scale.

In the case of the Doughnut, respondents whose work is related to policy making give the framework a lower average compatibility score (5.17), than those who do not work on policy making. Respondents who are focused on the city scale give the Doughnut an average score of 5.52, which is lower than those who are not concerned with the city scale (8.33).

Regarding Broad Prosperity, there is a difference between the average scores given by those who work on the regional scale (7.67) and those who do not work on the regional scale (6.00).

Table 4
Statistically significant Mann-Whitney U test results

Framework	Variable	U	Groups	Group mean
SDGs	Policy making	70.000*	No	6.37
			Yes	5.25
	City scale	30.500**	No	7.5
			Yes	5.56
	International scale	47.500*	No	5.58
			Yes	7.14
DE	Policy making	70.500*	No	6.63
			Yes	5.17
	City scale	18.000***	No	8.33
			Yes	5.52
BP	Beyond the city, within the Netherlands scale	56.500**	No	6.00
			Yes	7.67

*Note: N=31 for all three frameworks. *p < .10, **p < .05, ***p < .01.*

4.2 The explanatory power of civil servants' position in the organization on their evaluation of the frameworks

Table 5 presents the result of correlation analysis between the scores of the three frameworks on a scale of 1 to 10 and the organizational cluster that civil servants work at (whether or not they work at Ruimte & Economie), their work's relationship to policy and the geographical focus of their work. Spearman's ρ correlation coefficient was used and a significance level of $p < 0.10$ was accepted. Statistically significant correlations are highlighted in bold. Tables 6, 7 and 8 present the regression models explaining the compatibility scores of the SDGs, Doughnut Economics and Broad Prosperity on a 1 to 10 point scale. In the paragraphs below the regression models are described per framework.

Table 5
Spearman’s rho correlations between framework scores on a 1 to 10 scale and predictor variables.

		1. SDGs score	2. DE score	3. BP score	4.	5.	6.	7.	8.	9.	10.
Frameworks	1. SDGs score										
	2. DE score	**.633**									
	3. BP score	**.468**	**.410**								
Cluster	4. R&E	-.041	.009	-.208							
Work’s relation to policy	5. Policy making	**-.331**	**-.325**	.117	.016						
	6. Policy implementation	.190	.207	-.116	-.052	-.114					
	7. Policy evaluation	-.113	.037	.219	.173	**.592**	.054				
Geographical focus	8. < city	.283	.285	-.148	.193	-.077	**.376**	.056			
	9. City	**-.412**	**-.525**	.051	-.159	**.389**	-.033	.222	-.116		
	10. > city, within the Netherlands	.071	-.239	**.434**	-.297	**.320**	-.281	.048	-.210	**.389**	
	11. International	**.319**	.222	.202	.204	-.112	-.126	-.112	-.285	-.126	.204

Note: for correlations with Cluster R&E N=29, otherwise N=31. Bold correlations are significant (p < .10)

SDGs

Regression model 1 indicates that whether or not civil servants work at the cluster Ruimte & Economie does not explain the compatibility score they give to the SDGs (Adjusted $R^2 = -0.037$, $p > 0.10$). The second model, which assesses the effect of the relationship of civil servants' work to policy indicates that this variable is also not of influence on the compatibility score (Adjusted $R^2 = 0.080$, $p > 0.10$; R^2 Change = 0.074, $p > 0.10$). Model 3 indicates that the compatibility score of the SDGs is influenced by the geographical scale on which a civil servants' work is focused (Adjusted $R^2 = 0.228$, $p < 0.10$; R^2 Change = 0.374, $p < 0.05$). In particular, the regression coefficients of whether their work is related to the city scale or higher levels of national government are statistically significant. If civil servants indicate that their work is focused on the city scale, the model estimates that a lower score is given to the SDGs' compatibility. If civil servants indicate that their work is focused on higher levels of Dutch government, the model estimates that a higher compatibility score is given.

Table 6
Regression models SDGs compatibility scores

	Model 1		Model 2		Model 3	
	β	SE	β	SE	β	SE
Cluster						
Ruimte & Economie	-.012	.821	-.007	.861	-.079	.836
Work's relation to policy						
Policy making			-.317	.981	-.270	.902
Policy implementation			.039	1.189	.080	1.052
Policy evaluation			.149	1.042	.207	.898
Geographical focus						
< city					.270	.775
City					-.423**	.950
> city, within the Netherlands					.386*	.884
International					.291	.957
Adjusted R^2	-.037		-.080		.228*	
Change in R^2	.000		.074		.374**	
F-value	.004		.481		2.035	
Residual degrees of freedom	27		24		20	

*Note: *p < .10, **p < .05. N=29*

Doughnut Economics

Regression model 1 indicates that whether or not civil servants work at the cluster Ruimte & Economie does not explain the compatibility score they give to the Doughnut (Adjusted $R^2 = -0.037$, $p > 0.10$). The second model, which additionally includes civil servants' work' relationship to policy as a predictor, is not statistically significant (Adjusted $R^2 = 0.069$, $p > 0.10$; R^2 Change = 0.202, $p > 0.10$). Although, two regression coefficients are statistically significant at $p < 0.05$: the regression coefficients for whether civil servants' work is related to policy making and whether their work is related to policy evaluation. If civil servants indicate that their work is related to policy making, the model estimates that a lower score is given to the Doughnut's compatibility. If their work is focused on policy evaluation, the compatibility score is higher. Model 3 indicates that the compatibility score of the Doughnut is influenced by the geographical scale on which a civil servants' work is focused (Adjusted $R^2 = 0.294$, $p < 0.05$; R^2 Change = 0.294, $p < 0.05$). In particular, the regression coefficient of whether their work is related to the city scale is statistically significant. If civil servants indicate that their work is focused on the city scale, the model estimates that a lower score is given to the Doughnut's compatibility.

Table 7
Regression models DE compatibility scores

	Model 1		Model 2		Model 3	
	β	SE	β	SE	β	SE
Cluster						
Ruimte & Economie	-.019	.866	-.096	.843	-.268	.844
Work's relation to policy						
Policy making			-.490**	.961	-.299	.910
Policy implementation			-.052	1.165	-.049	1.061
Policy evaluation			.480**	1.020	.481**	.907
Geographical focus						
< city					.224	.782
City					-.487**	.959
> city, within the Netherlands					.023	.892
International					.269	.965
Adjusted R^2	-.037		.069		.294**	
Change in R^2	.000		.202		.294**	
F-value	.009		1.519		2.460**	
Residual degrees of freedom	27		24		20	

*Note: *p < .10, **p < .05. N=29*

Broad Prosperity

In the case of Broad Prosperity, none of the models using cluster, work's relation to policy and the geographical focus of civil servants as predictors are significant. Only in model 3, the regression coefficient of whether civil servants' work is focused on the geographical scale beyond the city, but within the country is significant at $p < 0.10$. It indicates that when civil servants' work is focused on this particular geographical scale, that they give a higher compatibility score to Broad Prosperity than if their work is not concerned with this scale.

Table 8
Regression models BP compatibility scores

	Model 1		Model 2		Model 3	
	β	SE	β	SE	β	SE
Cluster						
Ruimte & Economie	-.234	.777	-.281	.824	-.251	.893
Work's relation to policy						
Policy making			-.084	.939	-.136	.963
Policy implementation			-.024	1.138	.066	1.124
Policy evaluation			.270	.997	.346	.960
Geographical focus						
< city					.034	.828
City					-.215	1.015
> city, within the Netherlands					.411*	.945
International					.217	1.022
Adjusted R^2	.020		-.044		.072	
Change in R^2	.055		.050		.232	
F-value	1.564		.705		1.271	
Residual degrees of freedom	27		24		20	

*Note: *p < .10, **p < .05. N=29*

4.3 The frameworks' compatibility with five work dimensions

Table 9
Average score of the compatibility of the SDGs, the Doughnut and Broad Prosperity with five dimensions of civil servants' work measured on a 3-point Likert Scale.

	SDGs		The Doughnut		Broad Prosperity		
	Mean	SD	Mean	SD	Mean	SD	X^2
Goal formulation	1.94	.629	1.97	.795	2.23	.560	3.284
Providing insights into considerations	1.74	.682	1.90	.700	2.29	.643	11.861***
Evaluating results	1.90	.746	1.90	.700	1.97	.657	.348
Internal communication	1.87	.763	1.74	.729	2.23	.717	8.029**
External communication	2.03	.752	1.74	.682	2.16	.779	9.270**

*Note: N = 31, Min = 1, Max = 3 for all. *p < .10, **p < .05, ***p < .01*

Table 9 presents how civil servants scored the frameworks' compatibility with five dimensions of their work on a 3-point Likert scale. The figures in Appendix C depict the distributions of responses across the three potential scores. The outcome of Friedman tests performed for each of the frameworks to assess whether there are significant differences between how respondents evaluate the dimensions within each of the frameworks is that there are no significant differences (differences between values vertically in table 9). The outcome of Friedman tests that assess whether there are significant differences between framework evaluations within the same dimensions indicate that there are significant differences between framework evaluations within the dimensions of providing insight into considerations ($X^2 = 11.861$, $p < 0.01$), internal communication ($X^2 = 8.029$, $p < 0.05$) and external communication ($X^2 = 9.270$, $p < 0.05$).

Post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in an accepted significance level of $p < 0.033$. Within the dimension of 'providing insight into considerations', there is no significant difference between the evaluations of the SDGs and the Doughnut ($Z = -1.387$, $p = 0.166$). There are significant differences between the evaluations of the SDGs and Broad Prosperity ($Z = -3.145$, $p = 0.002$) and the evaluations of the Doughnut and Broad Prosperity ($Z = -2.288$, $p = 0.022$).

In the dimension of internal communication there are no significant differences between the evaluations of the SDGs and the Doughnut ($Z = -0.821$, $p = 0.412$), nor between the SDGs and Broad Prosperity ($Z = -2.001$, $p = 0.045$). There is a significant difference between the evaluations of the Doughnut and Broad Prosperity ($Z = -2.696$, $p = 0.007$).

In the dimension of external communication there is no significant difference between the evaluations of the SDGs and the Doughnut ($Z = -2.066$, $p = 0.039$), nor between the SDGs and Broad Prosperity ($Z = -0.943$, $p = 0.346$). There is a significant difference between the evaluations of the Doughnut and Broad Prosperity ($Z = -2.829$, $p = 0.005$).

4.4 The role of five dimensions of compatibility in determining overall compatibility

Tables 10, 11 and 12 show the correlations between five dimensions of framework compatibility and overall compatibility scores given to the three frameworks on a 1 to 10 point scale. In the case of the SDGs, positive correlations are found between the evaluations for goal formulation, providing insight into considerations and internal communication on the one hand, and the overall compatibility score on the other hand ($p < 0.01$). Significant positive correlations also exist between the evaluations for providing insight into considerations, evaluation of results and internal communication on the one hand, and goal formulation on the other. There is a positive correlation between the frameworks' compatibility with providing insight into considerations and its compatibility with evaluating results. The final correlation in the case of the SDGs is a positive one between the framework's compatibility with internal communication and external communication.

Table 10
Correlations between framework compatibility evaluations of five dimensions and overall compatibility score of the SDGs using Spearman's rho correlation coefficient.

	1. SDGs score	2.	3.	4.	5.
1. SDGs score					
2. Goal formulation	**.514*** **				
3. Providing insight into considerations	**.539*** **	**.387** **			
4. Evaluation of results	.224	**.498*** **	**.432** **		
5. Internal communication	**.474*** **	**.347* **	.252	.056	
6. External communication	.243	.144	.043	.238	**.587*** **

*Note: bold correlations are significant. *p < .10, **p < .05, ***p < .01*

In the case of the Doughnut, positive correlations are found between the overall compatibility score one the one hand, and all dimensions except for external communication on the other hand (p < 0.01). Additional correlations came forward between goal formulation and providing insight into considerations, evaluation of results and internal communication. The degree of compatibility of the Doughnut with providing insight into considerations is positively correlated with its compatibility with the evaluation of results. Also, the framework's evaluated compatibility for internal communication is positively correlated with its compatibility with external communication.

Table 11
Correlations between framework compatibility evaluations of five dimensions and overall compatibility score of the Doughnut using Spearman's rho correlation coefficient.

	1. DE score	2.	3.	4.	5.
1. DE score					
2. Goal formulation	**.594*** **				
3. Providing insight into considerations	**.460*** **	**.478*** **			
4. Evaluation of results	**.646*** **	**.612*** **	**.606*** **		
5. Internal communication	**.502*** **	**.679*** **	.268	**.462*** **	
6. External communication	.148	.195	.277	**.349* **	**.502*** **

*Note: bold correlations are significant. *p < .10, **p < .05, ***p < .01*

In the case of Broad Prosperity, the overall compatibility score is significantly positively correlated with all dimensions of compatibility (p < .10). The compatibility evaluation of goal formulation is positively correlated with providing insight into considerations, internal communication and external communication. Moreover, The evaluation of compatibility with providing insights into considerations is positively correlated with compatibility with evaluating results. Lastly, there is a positive correlation between the compatibility evaluation of Broad Prosperity for internal communications with its compatibility for external communication.

Table 12
Correlations between framework compatibility evaluations of five dimensions and overall compatibility score of Broad Prosperity using Speaman's rho correlation coefficient

	1. BP score	2.	3.	4.	5.
1. BP score					
2. Goal formulation	**.324***				
3. Providing insight into considerations	**.339***	**.393****			
4. Evaluation of results	**.501*****	.264	**.727*****		
5. Internal communication	**.419****	**.500*****	.099	.155	
6. External communication	**.337***	**.504*****	-.030	.128	**.600*****

*Note: bold correlations are significant. *p < .10, **p < .05, ***p < .01*

Table 13 presents the results of the multiple linear regressions that assess the extent to which the scores on five dimensions of compatibility with work explain the overall compatibility score given to the three frameworks. The models assessing the influence of the evaluations of five dimensions of framework compatibility can account for 38.8% of the variance in SDGs compatibility scores. Out of the individual dimensions, only the coefficient of the dimension 'providing insights into considerations' is statistically significant at $p < 0.10$ and indicates that a higher evaluation of this dimension is accompanied by a higher SDGs score on a 1 to 10 point scale. Only the coefficient for the dimension 'evaluation of results' has a negative value in the model.

Table 13
Regression models explaining the frameworks' overall compatibility scores by the scores of compatibility with four work dimensions.

	Model 1: SDGs		Model 2: DE		Model 3: BP	
	β	SE	β	SE	β	SE
Goal formulation	.307	.573	.165	.619	.122	.814
Providing insight into considerations	.313*	.504	.136	.559	.083	.825
Evaluating results	-.039	.503	.417	.610	.321	.720
Internal communication	.299	.538	.258	.675	.220	.567
External communication	.076	.521	-.179	.565	.087	.547
Adjusted R^2	.388***		.442***		.221**	
F-value	4.796***		5.750		2.697	
Residual degrees of freedom	25		25		25	

*Note: *p < .10, **p < .05., ***p < .01. N=31*

In the cases of the Doughnut and Broad Prosperity the regressions resulted in models with significant, substantial Adjusted R^2 values, but no individually significant predictors. The model assessing the influence of the evaluations of five dimensions of framework compatibility can account for 44.2% of the variance in Doughnut compatibility scores. The model assessing the influence of the evaluations of five dimensions of framework compatibility can account for 22.1% of the variance in Broad Prosperity compatibility scores. None of the VIF values exceed 10 nor do any of the other collinearity diagnostics point in a clear direction for collinearity. See Appendix D for the complete collinearity diagnostics tables.

S4.5 Civil servants' attitudes towards the use of single or multiple frameworks

A total of 49 respondents replied 'disagree', 'neutral' or 'agree' to all three of the statements. To statement 1 "*It is best to use a single framework for the entire process of policy making, implementation and evaluation*" 40% of respondents agree, while around 30% disagree. A little over 45% agree that the right framework depends on the policy theme with over 35% disagreeing. A little over 40% find it problematic that different departments use different frameworks, while about 35% do not find it problematic.

A correlation analysis shows that there are statistically significant relationships between respondents' attitudes towards the three statements (table 14). If civil servants have a more positive attitude towards using one framework for the entire policy process (1), they have a more negative attitude towards the idea that the right framework depends on the policy theme (2) ($p < 0.05$). If civil servants have a more positive attitude towards using one framework for the entire policy process (1), they find it more problematic that different departments use different frameworks (3) ($p < 0.01$). There is no significant correlation between civil servants' attitudes towards statements two and three.

Table 14
Spearman's rho correlations between three statements about the use of single or multiple frameworks.

	1.	2.
1. It is best to use a single framework for the entire process of policy making, implementation and evaluation.		
2. The appropriate framework depends on the policy theme.	**-.362****	
3. It is problematic that different departments use different frameworks.	**.571*****	-.236

*Note: bold correlations are significant. **p < .05, ***p < .01*

In the contingency table of statements 1 and 2 (table 15) respondents who lean towards the use of single frameworks can be found in the bottom left corner, adding up to almost 39% of respondents. Respondents who lean towards the use of multiple frameworks are found in the upper right corner, adding up to almost 43% of respondents. 4% are completely neutral and the rest are inconsistent in their attitudes towards single or multiple frameworks.

Table 15
Crosstabulation statements 1 and 2.

		Statement 2			
		Disagree	Neutral	Agree	Total
Statement 1	Disagree	3	4	10	15
	Neutral	3	2	7	14
	Agree	12	4	6	20
Total		18	8	23	49

In the contingency table of statements 1 and 3 (table 16) respondents who lean towards the use of single frameworks can be found in the bottom right corner, adding up to almost 43% of respondents. Respondents who lean towards the use of multiple frameworks are found in the upper left corner, adding up to almost 33% of respondents. 14% are completely neutral and the rest are inconsistent in their attitudes towards single or multiple frameworks.

Table 16
Crosstabulation statements 1 and 3.

		Statement 3			
		Disagree	Neutral	Agree	Total
Statement 1	Disagree	11	1	3	15
	Neutral	4	7	3	14
	Agree	2	3	15	20
Total		17	11	21	49

5. Discussion

5.1 Interpretation of the Results

5.1.1 Beyond GDP frameworks in the minds of civil servants

For the purpose of analysis and for the researcher and the reader to be on the same page, an understanding of beyond GDP frameworks was set out at the start of this research. However, it is unsure whether that same definition is shared by the civil servants who participated in the study. One interviewee who works on the development of Broad Prosperity shared the perception that all three of the frameworks are aiming for practically the same goals. To them it does not matter under which framework name those goals are achieved. Thus according to them, and likely others as well, the frameworks can be considered interchangeable in terms of the goals they aim for and not much more than a new name. On the other hand, another interviewee emphasized that conceptualizations and philosophies direct conversations, decisions and outcomes that follow from the use of a particular framework. Therefore, it is possible that some civil servants in reality evaluated only part(s) of the frameworks, such as the indicators or targets, ignorant of their underlying philosophies and fundamental conceptualizations. As important as it is to set clear theoretical boundaries, the understandings of research participants also matter. Since no shared understanding of beyond GDP frameworks was established between the researcher and participants, the research results must be interpreted keeping in mind that they follow from whatever understanding of the frameworks civil servants had.

5.1.2 The role of civil servants' position in the organization for the evaluation of the overall compatibility of the SDGs, the Doughnut and Broad Prosperity

Due to the small sample and its composition, the variable 'cluster' could only account for two groups: civil servants who are part of the cluster Ruimte & Economie and those who are not. Based on this availability of data, no differences between these groups were found with the Mann-Whitney U test nor was a relationship found between the cluster that civil servants work at and their evaluation of the overall compatibility of the SDGs, the Doughnut and Broad Prosperity. The other clusters could not be accounted for separately, so nothing can be said about whether the evaluations of civil servants at the other four department differ from each other and Ruimte & Economie. Moreover, in reality the clusters together are made up of 45 departments, each with their own thematic orientation. Considering the great diversity still within clusters, it is possible that differences based on thematic orientation do exist, but that they cannot be captured at the level of clusters.

While statistical differences were found between the overall compatibility of the frameworks given by groups of civil servants based on their relation to policy and their geographical focus, the extent to which these differences can be explained by their position in the organization varies. Work's relation to policy holds explanatory power for the overall compatibility score of only the Doughnut and geographical focus holds explanatory power for the overall compatibility scores of all three of the frameworks. Whether civil servants' work is related to policy evaluation has a significant positive influence on the overall compatibility score they assign to the Doughnut, which is not the case for the SDGs and Broad Prosperity.

In the cases of the SDGs and the Doughnut, the frameworks are found less compatible by civil servants if their work is focused on the city scale. This is not the case for Broad Prosperity. A potential explanation for this can be found in the design of the frameworks and the intent behind their use. The SDGs and the Doughnut have been designed with the global scale in mind as a starting point to develop goals and accompanying indicators to measure performance towards their achievement. Previous studies have found that such frameworks and their indicators have to somehow be localized (Ciambra et al., 2023) or downscaled (Turner & Wills, 2022) to be applicable to the city scale. Broad Prosperity on the other hand, is a monitoring framework designed from a principle that remains the same regardless of geographical scale and with built in space to include themes that are more relevant to lower scales (Horlings & Smits, 2019). As such, it is not significantly less compatible with use on the city scale.

Broad Prosperity and the SDGs are considered more compatible if civil servants' work focuses on geographical scales beyond the city, but within the Netherlands, which does not apply to the Doughnut. In the case of Broad Prosperity, this is possibly due to the fact that the framework is already used by the national and regional governments. Multiple interviewees mentioned that national and regional budgets have been made on the basis of a Broad Prosperity approach and that some subsidies are given accordingly. While the Doughnut was originally designed at the global scale, there was a recognition that action at the city scale would be necessary to achieve the goals set out and Amsterdam became one of the pioneering cities for the development of localized Doughnut Economics frameworks. In this sense, the development of the framework for the city specifically may not make the framework more compatible with working on higher geographical scales within the country, where the framework is not being used to the same extent.

5.1.3 The SDGs, the Doughnut and Broad Prosperity and five dimensions of civil servants' work

Despite varying average scores for the compatibility of the frameworks with five dimensions of civil servants' work, none of the frameworks are significantly more or less compatible with one or more of them. This means, for example, that the SDGs are not considered significantly more compatible with external communication than they are with providing insight into considerations. The same logic applies to the evaluations of the Doughnut and Broad Prosperity with the five dimensions. However, within the dimensions of providing insight into considerations, internal communication and external communication, there are significant differences between how the three frameworks score. At the same time there were no significant differences between the overall scores of the frameworks. This indicates that compatibility differences between the frameworks become apparent when assessing more specific needs that they are required to fulfill.

Broad Prosperity is considered more compatible with providing insight into considerations than both the SDGs and the Doughnut. In the dimensions of internal and external communication, Broad Prosperity significantly outperforms the Doughnut, but not the SDGs. Interviewees mentioned that the SDGs' lack of clear connections between the goals and the Doughnut's high level of abstraction as problematic. Broad Prosperity as a consideration framework should help overcome those hurdles. A possible explanation for its greater compatibility is that the group within the municipality that is responsible for the development of this Broad Prosperity framework has also developed a number of instruments, like a serious game, that directly link the rationale of the framework to other civil servants' work. This does beg the question whether it is the conceptual framework and its operationalization as such that is considered more compatible or the thought process that is outlined and instruments that have been developed to translate its rationale from the more abstract realm to the practical. Moreover, upon the question why the choice was not instead made to develop an 'SDGs consideration framework' or 'Doughnut consideration framework', the response was that 1) it can be considered as such and 2) that is was considered beneficial to align with the Broad Prosperity use of higher levels of Dutch government.

Further analysis of the extent to which compatibility with the five dimensions explains the overall compatibility of the frameworks shows that the extent to which they are influential differs per framework. The compatibility with the five chosen dimensions explains around half of the variance in overall Broad Prosperity compatibility scores compared to the SDGs and the Doughnut. Especially for Broad Prosperity, this suggests that there are other dimensions of civil servants' work that beyond GDP frameworks may be more or less compatible with. Moreover, while the models have quite high R^2 values, only the model of the SDGs contains a single significant regression coefficient. While the VIF values were well below 10, the outcomes still suggest that there are issues with collinearity.

5.1.4 Single or multiple frameworks: what does it matter?

Opinions about the use of single or multiple frameworks throughout the organization differ, but there is some alignment between respondents' answers to the statements. There is a tendency that those who prefer a single framework for the entire policy process are less in agreement that the right framework depends on the theme at hand, and they also find it more problematic that different departments use different frameworks. However, the correlations are not very strong. A minority is either completely neutral towards the diversity of frameworks being used throughout the organization or has inconsistent attitudes.

Interviews provided some insight in the attitudes towards the use of single or multiple frameworks. While the use of multiple frameworks can be experienced as confusing by some, others argue that adhering to a single framework should not be the goal. Frameworks like the Sustainable Development Goals, Doughnut Economics and Broad Prosperity can all somehow offer a (slightly) different perspective through the rationales, dimensions and indicators that they include, while excluding others. Directing complete focus towards one framework puts those other valid perspectives out of scope. This can create blind spots similar to the siloes that the municipal organization knows nowadays and that more holistic frameworks argue should be broken. Moreover, depending on the actors involved, different frameworks may be useful for different situations, be it for gathering ideas according to themes or providing a structure to trade-off conversations.

5.2 Feedback to the literature

The focus of the use of beyond GDP frameworks as paradigm innovations from a theoretical standpoint is rather clear from the definition by Francis & Bessant (2005) and the definition of De Vries et al. (2016) of conceptual innovation helps us image what kinds of 'things' may be introduced in an organization: new concepts, frames of reference or paradigms. In the case of beyond GDP frameworks at the municipality of Amsterdam, there seems to be conceptual confusion, which is when different actors attach different meaning to concepts (Whitby et al., 2014). The fact that some civil servants at the city level who are actively involved in the framework developments believe that one framework can just be considered a more practical 'follow-up' framework of another, that others find that the frameworks aim for basically the same goals and that yet others recognize that the frameworks are distinct in their conceptualizations points in the direction of such conceptual confusion also exists at the municipality of Amsterdam. Since the introduction of new concepts is at the core of beyond GDP frameworks, it can be argued that there is confusion throughout the organization about *what* has been adopted and what is being developed and implemented in the first place.

Even when there is conceptual clarity, concepts, frames of reference and paradigms can still be considered rather abstract to city government practitioners – too abstract to be implemented for instrumental use as they are at a municipal organization. Not all beyond GDP frameworks as paradigm innovations come with a detailed instruction manual for their use, if any at all. However, within the municipality of Amsterdam Broad Prosperity is being developed explicitly as a framework for *considerations* and *ex-ante evaluation of policies*. With this use in mind, additional products were developed to establish the change in mindset in civil servants *and* bridge the gap between the change in mindset and a change in actions. This is in line with the suggestions of Hayden (2021) to achieve instrumental use of indicators. Buchheim et al. (2016) suppose that other types of innovation can follow paradigm innovation. While this research cannot definitively claim that setting a clear use intent and developing complementary innovations is the reason why the framework is considered more compatible than the other two in some instances, it is possible that the implementation of paradigm innovations benefits from complementary product and process innovations that improve their functionality. Thus, not only can other types of innovations follow from paradigm innovation, but other types of innovations may be necessary for successful implementation of paradigm innovations.

Implementation literature has looked mostly at single product and process innovations in the context of private organizations and it assumes that the implementation stage has been preceded by stages in which problems are identified and considerations are made about whether an innovation is the right fit to solve the problem after which adoption decisions are made. This is the approach to innovation that Korteland & Bekkers (2008) describe as the logic of functionality. Due to the context being a public organization, not only a logic of functionality but also a logic of appropriateness plays a role. As such, beyond GDP frameworks might be considered compatible with political use by political actors following from the logic of appropriateness, but less so instrumentally following from a logic of functionality that is relevant for civil servants. Add to this conceptual confusion about what is being implemented for either purpose and it can be supposed the adoption and implementation of paradigm innovations at local governments follows different trajectories than those that are traditionally outlined in the literature. Their trajectories may result in the adoption and need for implementation of paradigm innovations that are neither fully understood, and thus potentially conflicting, nor considered functionally compatible for the work that civil servants are required to perform.

5.3 Reflection on the Methods

5.3.1 Survey Dissemination: snowballing and the structure of the organization

The snowball sampling method in combination with structure of the organization of the municipality proved to be challenging for reaching the desired quantity and diversity of respondents. While a pre-selection of who was included in the study was not desired, selection was in the end made by civil servants themselves. The researcher was dependent on civil servants to distribute the survey within their networks. Some of them asked which of their colleagues were considered part of the target group of the survey, to which the response was: everyone. Still, based on their assumptions about the relevance of the topic for their colleagues, their colleagues' knowledge on the topic or their colleagues' interest in the topic of sustainability, decisions were made about who to forward the survey to and thereby who to exclude. Multiple people mentioned that they wanted to be considerate of their colleagues' time by only forwarding the survey if it was really applicable to them.

The difficulty of disseminating the survey was shared with the research and statistics department of the municipality of Amsterdam (Onderzoek, Informatie & Statistiek, OIS). OIS admitted that they themselves are hesitant about distributing surveys due to the large amount of emails that civil servants already receive and that they also struggle with gathering responses to their own surveys. They said that it is very hard to get responses from civil servants if the topic does not influence their daily job. This may have been a factor at play in the case of this study. The current structure of the municipality is thematically separated, also described as siloed, contrary to the more integrated/holistic approach advocated for through the frameworks. As a consequence, those who are working within such siloes sometimes do not recognize the relevance of the topic of the study for themselves and/or others. This can reinforce the sentiment that the survey is not relevant for themselves and/or their colleagues. The first recommendation from OIS was to work with more specified target audiences based on department, program or (lack of) knowledge about the frameworks, for example. While this may indeed be effective for gathering a larger number responses, it is worth considering whether the sampling outcome of that approach is in line with the principle that targeted users of the frameworks are not limited to specific types of divisions in the organization.

Another recommendation was to either conduct surveys in person or to organize group discussions with pre-selected civil servants. They acknowledged that such group discussions could provide strong qualitative results rather than quantitative. While qualitative results can be scientifically strong, the extent to which they are considered 'worthy' to be used in practice depends on the value that the municipal actors assign to qualitative outcomes compared to quantitative outcomes.

5.3.2 Survey Design

Response options

The survey was designed with the reality in mind that there is variety in the number of beyond GDP frameworks that civil servants work with, if any at all. As a consequence, the amount of data is not equal per relevant variable. In order to make comparisons between frameworks, the choice was made to only include respondents who had scored all three of the frameworks. In the case of this study with an already small sample, this did mean that a relatively large amount of data had to be discarded. Even if the study were to be performed on a larger scale with potential for generalization beyond the sample, this would still only extend to civil servants with the capacity to score all three frameworks. Since the frameworks are not structurally embedded throughout the organization, this can limit the extent to which the whole organization can be accounted for.

Survey compatibility with target groups

A number of respondents commented that they would have preferred to have more information provided about the frameworks in the survey. A deliberate decision was made not to include any details about the frameworks in the survey due to its estimated duration and effort required from respondents. To fully capture the three frameworks in all their dimensions and connect them to ones daily tasks was considered unfeasible in an exercise of 10 to 15 minutes. This does raise the question whether the chosen method of data collection, a survey, and its design are appropriate for all groups throughout the municipality who are expected to be affected by the use of one or more of the three frameworks. Maybe the survey should be adjusted to the different target groups based on their knowledge about the frameworks or the survey as it has been designed should be distributed to civil servants with sufficient knowledge about them.

6. Conclusion

To summarize, this thesis stems from the problem that beyond GDP approaches have trouble getting integrated into practice. This is however necessary in order to replace the 'economic-growth-GDP-compass' guiding government decision-making and to maximize their potential impact. The aim of the research was to gain insight into the user perspective on the implementation of beyond GDP frameworks in the city government context to inform the development of beyond GDP frameworks that have a higher potential for implementation to achieve deep system change. The leading research question was how civil servants of the municipality of Amsterdam evaluate three beyond GDP frameworks for implementation. The three frameworks are the Sustainable Development Goals, Doughnut Economics and Broad Prosperity. The frameworks were conceptualized as paradigm innovations that aim for change in the realm of intent of a system, in this case the organization of the municipality of Amsterdam. Out of all the factors that influence the implementation of innovation, the focus was placed on the compatibility of the frameworks with the work of civil servants according to themselves.

Overall, civil servants find the Sustainable Development Goals, Doughnut Economics and Broad Prosperity equally, slightly compatible with their work. There are some differences between how subgroups of civil servants evaluate each of the frameworks depending on their work's relation to policy and the geographical scale that they work on. In the case of the SDGs there are differences depending on whether civil servants' work relates to policy making, whether their work is focused on the city scale and whether it is focused on the international scale. In the case of the Doughnut, there are differences depending on whether civil servants work relates to policy making and whether their work is focused on the city scale. In the case of Broad Prosperity there are differences depending on whether they work on the scale beyond the city, but within the country. Based on the research performed, it remains unknown which aspects of the frameworks were (not) considered in the minds of civil servants when they made the evaluations, but it is not unlikely that a degree of conceptual confusion underlies the outcomes.

The extent to which civil servants' position in the organization actually explains the overall compatibility scores they assign to the frameworks differs per framework. Geographical scale is a significant explanatory factor in the case of the SDGs and the Doughnut. In particular, if civil servants work on the city scale they find the frameworks less compatible. If civil servants' work is concerned with the scale beyond the city, but within the country, they find the SDGs more compatible. Their work's relation to policy has significant explanatory power in the case of overall compatibility of the Doughnut, in which case the framework is considered more compatible by policy evaluators. Thus, some but not all differences between overall compatibility scores throughout the organization can be attributed to civil servants' position in the organization.

When looking per framework how civil servants evaluated their compatibility with five dimensions of their work, none of the frameworks excel in any dimension compared to the other dimensions. However, there are significant differences between how the frameworks' compatibility is evaluated within the dimensions of providing insight into considerations, internal communication and external communication. Broad Prosperity is considered more compatible with providing insight into considerations than the other two frameworks and Broad Prosperity scores significantly better than the Doughnut in the other two dimensions. The question remains whether this can be fully attributed to the Broad Prosperity framework itself or whether its complementary instruments play a role.

While compatibility with these five dimensions of civil servants' work explains significant amounts of the variance in their overall compatibility, they explain only half of the variance of Broad Prosperity compatibility compared to the SDGs and Doughnut Economics. Due to hidden collinearity in the regression models, the explanatory power of compatibility with each of the individual dimensions remains unclear.

Within the municipality of Amsterdam the majority of civil servants are either consistently in favor of the use of single or multiple frameworks, divided over similarly sized groups.

6.1Recommendations for practice

The recommendations for practice are targeted at those who are involved in beyond GDP framework development and implementation for the municipality of Amsterdam. The first recommendation is to focus on establishing conceptual understanding of a range of beyond GDP frameworks amongst key players in the organization, including political actors, so they can make informed decisions regarding their adoption and implementation. It is problematic if conceptual confusion, and thus confusion about the very core of what beyond GDP frameworks are, continues to underlie their further development and implementation, and thus their potential impact. The second recommendation is to be critical towards the appropriate use and targeted users of (a multitude of) chosen beyond GDP frameworks with the goal of deep system change in mind. In order to achieve appropriate and committed use of an innovation by targeted users, those two elements have to be defined. Different use types and users may come forward that can benefit from particular instruments that make an alternative framework of thought actionable in their work context.

6.2 Future research

To close this report, two directions are suggested for future research. The first suggested research direction is to examine the pre-implementation trajectory of beyond GDP frameworks through the logics of functionality and appropriateness and how this may influence their implementation process. This can contribute to greater understanding of paradigm innovation compared the more-researched product and process innovations. Such research may be more qualitative in nature to uncover hidden assumptions and conceptual confusions that stand in the way of achieving deep system change through the use of beyond GDP frameworks.

The second direction lies in line with the research that has been performed here and is more quantitative in nature: to perform a similar study on a larger scale within the municipality of Amsterdam to explore how civil servants evaluate beyond GDP frameworks' compatibility depending on their position in the organization. In order for the researcher and civil servants to be on the same page regarding their understanding of what constitutes the frameworks under evaluation, explanatory sections should be added throughout the survey about what parts of the frameworks respondents are expected to evaluate. To get a sample that sufficiently covers the great diversity of groups, dissemination strategies that target members of specific teams more directly and setting quotas can be recommended. Other factors that may explain compatibility, like organizational values, may be included as predictors.

7. References

References

Aarons, G. A., Hurlburt, M. S., & Horwitz, S. M. (2010). Advancing a Conceptual Model of Evidence-Based Practice Implementation in Public Service Sectors. *Administration And Policy in Mental Health And Mental Health Services Research*, *38*(1), 4–23. https://doi.org/10.1007/s10488-010-0327-7

Abson, D. J., Fischer, J., Leventon, J., Newig, J., Schomerus, T., Vilsmaier, U., Von Wehrden, H., Abernethy, P., Ives, C. D., Jager, N. W., & Lang, D. J. (2016). Leverage points for sustainability transformation. *AMBIO*, *46*(1), 30–39. https://doi.org/10.1007/s13280-016-0800-y

Amsterdam donut coalitie. (n.d.). amsterdamdonutcoalitie.nl. https://amsterdamdonutcoalitie.nl/tools-en-tips (accessed 19 November 2024)

Baregheh, A., Rowley, J., & Sambrook, S. (2009). Towards a multidisciplinary definition of innovation. *Management Decision*, *47*(8), 1323–1339. https://doi.org/10.1108/00251740910984578

Baregheh, A., Rowley, J., & Hemsworth, D. (2016). The effect of organisational size and age on position and paradigm innovation. *Journal Of Small Business And Enterprise Development*, *23*(3), 768–789. https://doi.org/10.1108/jsbed-06-2015-0065

Bauer, M. S., & Kirchner, J. (2020). Implementation science: What is it and why should I care? *Psychiatry Research*, *283*, 112376. https://doi.org/10.1016/j.psychres.2019.04.025

Bleys, B. (2012). Beyond GDP: Classifying Alternative Measures for Progress. *Soc Indic Res* 109, 355–376. https://doi.org/10.1007/s11205-011-9906-6

Boarini, R., & D'Ercole, M. M. (2013). Going beyond GDP: An OECD Perspective*. *Fiscal Studies*, *34*(3), 289–314. https://doi.org/10.1111/j.1475-5890.2013.12007.x

Buchheim, L., Krieger, A., & Arndt, S. (2019). Innovation types in public sector organizations: a systematic review of the literature. *Management Review Quarterly*, *70*(4), 509–533. https://doi.org/10.1007/s11301-019-00174-5

Cavalletti, B., Corsi, M. (2018). "Beyond GDP" Effects on National Subjective Well-Being of OECD Countries. *Soc Indic Res* 136, 931–966. https://doi.org/10.1007/s11205-016-1477-0

Ciambra, A., Siragusa, A., Proietti, P., & Stamos, I. (2023). Monitoring SDG localisation: an evidence-based approach to standardised monitoring frameworks. *Journal of Urban Ecology*, *9*(1). https://doi.org/10.1093/jue/juad013

Damanpour, F., & Schneider, M. (2006). Phases of the Adoption of Innovation in Organizations: Effects of Environment, Organization and Top Managers1. *British Journal Of Management*, *17*(3), 215–236. https://doi.org/10.1111/j.1467-8551.2006.00498.x

Damanpour, F., & Schneider, M. (2008). Characteristics of Innovation and Innovation Adoption in Public Organizations: Assessing the Role of Managers. *Journal Of Public Administration Research And Theory*, *19*(3), 495–522. https://doi.org/10.1093/jopart/mun021

De Vries, H., Bekkers, V., & Tummers, L. (2015). INNOVATION IN THE PUBLIC SECTOR: A SYSTEMATIC REVIEW AND FUTURE RESEARCH AGENDA. *Public Administration*, *94*(1), 146–166. https://doi.org/10.1111/padm.12209

Francis, D., & Bessant, J. (2005). Targeting innovation and implications for capability development. *Technovation*, *25*(3), 171–183. https://doi.org/10.1016/j.technovation.2004.03.004

Geels, F. W., Elzen, B., & Green, K. (2004). General Introduction: System Innovation and Transitions to Sustainability. In *Edward Elgar Publishing eBooks*. https://doi.org/10.4337/9781845423421.00010

Gemeente Amsterdam. (2020). *Amsterdam Circulair 2020-2025*. https://www.amsterdam.nl/bestuur-organisatie/volg-beleid/duurzaamheid/circulaire-economie/ (accessed 19 November 2024)

Gemeente Amsterdam. (2021). *De Staat van de Stad Amsterdam XI*. https://onderzoek.amsterdam.nl/publicatie/de-staat-van-de-stad-amsterdam-xi-2020-2021 (accessed 19 November 2024)

Giovannini, E., & Rondinella, T. (2018). Going beyond GDP: theoretical approaches. In C. D'Ambrosio, *Handbook of Research on Economic and Social Well-Being* (pp. 1–51). Edward Elgar Publishing Limited.

Hayden, A. (2021). From Fantasy to Transformation: Steps in the Policy Use of "Beyond-GDP" Indicators. In *Springer eBooks* (pp. 119–139). https://doi.org/10.1007/978-3-030-67860-9_7

Horlings, E., & Smits, J. (2019). *Conceptueel kader voor een regionale monitor brede welvaart* (305185 ENR/AMR). CBS Den Haag. https://www.cbs.nl/-/media/_pdf/2019/41/regionale-monitor-brede-welvaart-2019.pdf (accessed 19 November 2024)

Implementation science. (z.d.). BioMed Central. https://implementationscience.biomedcentral.com/about

Jackson, T. (2016). Prosperity without Growth. In *Routledge eBooks*. https://doi.org/10.4324/9781315677453

Jansen, A., Wang, R., Behrens, P., & Hoekstra, R. (2024). Beyond GDP: a review and conceptual framework for measuring sustainable and inclusive wellbeing. *The Lancet Planetary Health*, *8*(9), e695–e705. https://doi.org/10.1016/s2542-5196(24)00147-5

Kim, J. S., & Chung, G. H. (2017). Implementing innovations within organizations: a systematic review and research agenda. *Innovation: Management, Policy And Practice*, *19*(3), 372–399. https://doi.org/10.1080/14479338.2017.1335943

Klein, K. J., & Sorra, J. (1996). The Challenge of Innovation Implementation. *Academy Of Management Review*, *21*(4), 1055–1080. https://doi.org/10.5465/amr.1996.9704071863

Korteland, E., & Bekkers, V. (2008). The diffusion of electronic service delivery innovations in dutch E-policing: The case of digital warning systems. *Public Management Review*, *10*(1), 71–88. https://doi.org/10.1080/14719030701763195

Kuznets S. National income, 1929–1932. https://fraser.stlouisfed.org/ title/national-income-1929-1932-971 (accessed 19 November 2024)

Meadows, D. (2015). Leverage points-places to intervene in a system. https://mchwdc.unc.edu/wp-content/uploads/2022/04/Leverage-Points_Places-to-Intervene-in-a-System-Meadows.pdf (accessed 19 November 2024)

Moore, M., & Hartley, J. (2008). Innovations in governance. *Public Management Review*, *10*(1), 3–20. https://doi.org/10.1080/14719030701763161

Raworth, K. (2018). *Doughnut economics: Seven ways to think like a 21st century economist*. Chelsea Green Publishing.

Real, K., & Poole, M. S. (2005). Innovation Implementation: Conceptualization and Measurement in Organizational Research. In *Research in Organizational Change and Development* (pp. 63–134). https://doi.org/10.1016/s0897-3016(04)15003-9

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F. S., Lambin, E., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J., Nykvist, B., de Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P. K., Costanza, R., Svedin, U., … Foley, J. (2009). Planetary Boundaries: Exploring the Safe Operating Space for Humanity. *Ecology and Society*, *14*(2). http://www.jstor.org/stable/26268316

Rogers, E. (2003) Diffusion of Innovations, 5th edn., New York: Free Press.

Seaford, C. (2013). The Multiple Uses of Subjective Well-Being Indicators. *Social Indicators Research*, *114*(1), 29–43. https://doi.org/10.1007/s11205-013-0381-0

Singh, S., Akbani, I., & Dhir, S. (2020). Service innovation implementation: a systematic review and research agenda. *Service Industries Journal*, *40*(7–8), 491–517. https://doi.org/10.1080/02642069.2020.1731477

Stiglitz, J.M., Sen, A. & Fitoussi, J. (2009). *Report by the commission on the measurement of economic performance and social progress*. Paris: Commission on the measurement of economic performance and social progress.

Turner, R. A., & Wills, J. (2022). Downscaling doughnut economics for sustainability governance. *Current Opinion in Environmental Sustainability*, *56*, 101180. https://doi.org/10.1016/j.cosust.2022.101180

United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development* (A/RES/70/1). https://undocs.org/en/A/RES/70/1 (accessed 19 November 2024)

UNECE/Eurostat/OECD (2014), Conference of European Statisticians Recommendations on Measuring Sustainable Development. New York/Genève: United Nations.

Whitby, A. (WFC) et al., 'BRAINPOoL Project Final Report: Beyond GDP - From Measurement to Politics and Policy' BRAINPOoL deliverable 5.2, A collaborative programme funded by the European Union's Seventh Programme for research, technological development and demonstration under grant agreement No. 283024. WFC (World Future Council), 31 March 2014.

Wisdom, J. P., Chor, K. H. B., Hoagwood, K., & Horwitz, S. M. (2013). Innovation Adoption: A Review of Theories and Constructs. *Administration And Policy in Mental Health And Mental Health Services Research*, *41*(4), 480–502. https://doi.org/10.1007/s10488-013-0486-4

Appendix A. Survey

WAGENINGEN
UNIVERSITY & RESEARCH

Introductie

. Welkom bij deze enquête van het project 'The Ideal(s) City: Monitoring what matters', een samenwerking tussen het AMS Institute en de Gemeente Amsterdam. Het project heeft als doel het ontwikkelen van een raamwerk om te meten wat de impact is van keuzes en projecten op de idealen van de stad.

De enquête gaat over hoe verschillende raamwerken aansluiten op het werk van ambtenaren. Bedankt dat je de tijd neemt om hem in te vullen!

De enquête bestaat uit vijf delen en het duurt ongeveer 10 minuten om hem in te vullen. Antwoorden worden anoniem verwerkt en alleen gebruikt voor onderzoeksdoeleinden binnen het AMS Institute.

Meer informatie of vragen? Neem een kijkje op de pagina van The Ideal(s) City project of stuur een bericht naar mary-ann.vankanten@wur.nl

Kennismaking

. In dit eerste deel zijn wij, van het projectteam, geïnteresseerd in het werk dat je doet als ambtenaar.

Q1. Stel je voor dat we tegenover elkaar zitten: hoe zou je het werk dat je doet in het kort beschrijven?

. De 20 beleidsthema's uit het coalitieakkoord

Afval en onderhoud	Bestuur en organisatie	Dieren	Diversiteit
Duurzaamheid	Economie	Gebiedsgericht werken	Groen
Innovatie	Internationale samenwerking	Kunst en cultuur	Onderwijs
Sport	Aanpak Binnenstad	Omgevingsvisie	Veiligheid
Verkeer en vervoer	Water	Werk en inkomen	Zorg en ondersteuning

Q2. Het huidige coalitieakkoord omvat de 20 beleidsthema's die hierboven zijn weergegeven. Aan welk(e) beleidsthema('s) is jouw werk vooral gerelateerd?

☐ ____________________

Q3. Hoe verhoudt het werk zich tot beleid? Het is mogelijk om meerdere antwoorden te selecteren.

- ☐ Beleid opstellen
- ☐ Uitvoering van beleid
- ☐ Evaluatie van beleid
- ☐ ____________________ Anders, namelijk

Q4. Op welk ruimtelijk schaalniveau is het werk gericht? Het is mogelijk om meerdere antwoorden te selecteren.

- ☐

Buurten en wijken

- ☐ Stadsdelen
- ☐ Gehele stad
- ☐ Regionaal
- ☐ Nationaal
- ☐ Internationaal

Maatschappelijke doelen

. In de tekstvakken hieronder staan drie beknopte uitspraken over een toekomst waar wij als maatschappij naar zouden kunnen streven.

Door met de muis over de tekst heen te bewegen verschijnt de optie om stukjes tekst te markeren.

Markeer in ieder tekstvak de delen die direct bij jouw werktaken passen in het **groen**, de delen die kunnen conflicteren met de werktaken in het **rood** en de delen die niet direct gerelateerd zijn aan het werk in het **geel**.

Q5. Uitspraak 1: Het doel waar we naar moeten streven is ...

het bestrijden van armoede in alle vormen en dimensies, het verminderen van ongelijkheid, het behouden van de planeet, het creëren van aanhoudende, inclusieve en duurzame economische en het bevorderen van sociale inclusie, voor de huidige generatie en toekomstige generaties.toekomstige generaties.

Q6. Uitspraak 2: Het doel waar we naar moeten streven is ...

een transitie naar een maatschappij die zich beweegt binnen de ruim door sociale tekortkomingen weg te werken en
binnen de grenzen van de natuur te blijven, zowel lokaal als mondiaal. De transitie moet gekenmerkt zijn door gendergelijkheid, sociale gelijkheid, politieke zeggenschap en vrede en gerechtigheid.gerechtigheid.

Q7. Uitspraak 3: Het doel waar we naar moeten streven is ...

het balanceren van de kwaliteit van leven in het hier en nu en de mate waarin deze ten koste gaat van die van latere generaties of van die van mensen elders in de wereld.van die van mensen elders

Q8. Op een schaal van 1 tot 10, hoe goed sluit iedere uitspraak over het algemeen aan bij jouw visie op de toekomst? 1 = sluit helemaal niet aan. 10 = sluit compleet aan.

	1	2	3	4	5	6	7	8	9	10
Uitspraak 1										
Uitspraak 2										
Uitspraak 3										

Q9. *Optioneel: wil je een toelichting geven op de beoordeling van de drie uitspraken?

Stellingen

Q10. Geef aan of je het (on)eens bent met de volgende vijf stellingen.

	Oneens	Neutraal	Eens	Weet ik niet
Economische groei is nodig als middel om menselijk welzijn te bevorderen.	○	○	○	○
Economische				

groei
kan
losgekop
peld
worden
van het
gebruik van

natuurlijke hulpbronnen door technologische ontwikkelingen.

Oneens Neutraal Eens Weet ik niet

Om
een
hoger
niveau
van
mense
lijk
welzij
n te
behal
en is
het
nodig
om de materiële
producti
e en
consump
tie te
beperken
tot
planetair
e
grenzen.

Sociale ongelijkheden kunnen alleen verminderd worden

door
het
bevor
dere
n van
inko

mens
groei.

Om sociale ongelijkheden te verminderen is een herverdeling van rijkdom tussen verschillende plekken op de wereld nodig.

De drie kaders

. Raamwerken bieden een structuur voor mensen en organisaties om de wereld om hen heen te bekijken en ermee om te gaan. Raamwerken geven richting aan wat wij als belangrijk zien en helpen ons aan te wijzen wat er goed gaat en wat mogelijk verbeterd kan worden.

De volgende vragen gaan over ervaringen met drie specifieke raamwerken:
Sustainable Development Goals, De (Stads)donut en Brede Welvaart.

Q11. Hoe vaak kom je de Sustainable Development Goals tegen in jouw werk?

- Nooit
- Zelden
- Af en toe
- Vaak

Q12. Hoe vaak kom je De (Stads)donut tegen in jouw werk?

- Nooit
- Zelden
-
- A
f
e
n
t
o
e
V

a

a

k

Q13. Hoe vaak kom je Brede Welvaart tegen in jouw werk?

- Nooit
- Zelden
- Af en toe
- Vaak

Q14. Op een schaal van 1 tot 10, hoe geschikt is ieder raamwerk om richting te geven binnen jouw werk?

Als je een raamwerk niet kent, selecteer dan de optie 'Raamwerk onbekend'.

	1	2	3	4	5	6	7	8	9	10	Raamwerk onbekend
Sustainable Development Goals											
De (Stads)donut											
Brede Welvaart											

Q15. Op basis van jouw ervaring, hoe geschikt is ieder raamwerk voor het vervullen van de volgende taken?

Als je een raamwerk niet kent, laat die kolom dan leeg.

1= Ongeschikt

2= Enigszins geschikt

3= Zeer geschikt

	Sustainable Development Goals	De (Stads)donut	Brede Welvaart

Het formuleren van

duidelijke doelen

Het inzichtelijk
maken van
afwegingen

Het evalueren van
resultaten

Communicatie
binnen de

Gemeente
Amsterdam

Externe communicatie (met

bewoners,
bedrijven,
andere
overheden, etc.)

Q16. *Optioneel: wil je een toelichting geven op ervaringen met het gebruik van één of meerdere raamwerken?

Verschillende kaders

Q17. Op dit moment worden er verschillende raamwerken gebruikt door de gemeente.

Geef aan of je het (on)eens bent met de volgende drie stellingen.

	Oneens	Neutraal	Eens	Weet ik niet
Voor het gehele proces van beleid vormen, uitvoeren, evalueren en erover communiceren kan het beste één en hetzelfde raamwerk worden gebruikt.				

Welk raamwerk
geschikt is, is

afhankelijk van het beleidsthema.

Oneens Neutraal Eens Weet ik niet

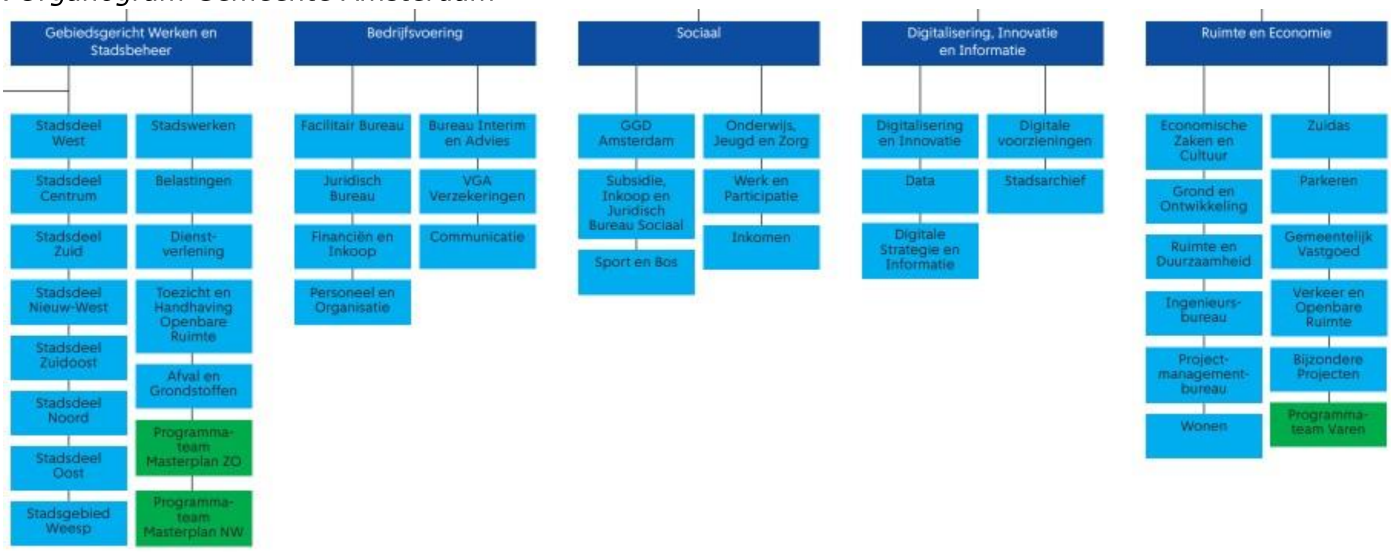
Het is
probl
emati
sch
dat
er verschil zit
tussen de

raamwerken die verschillende afdelingen gebruiken.

Afsluitende vragen

. Je hebt het laatste deel van de enquête bereikt! Deze laatste acht vragen gaan over een aantal achtergrondkenmerken.

. Organogram Gemeente Amsterdam



Q18. Bij welke afdeling ben je werkzaam?

Q19. Hoeveel jaren ben je werkzaam bij deze afdeling?

- <2
- 2-4
- 5-9
- 10-15
- >15

Q20. Hoeveel jaren ben je werkzaam bij de Gemeente Amsterdam?

- <2
- 2-4
- 5-9
- 10-15
- >15

Q21. In welk dienstverband werk je bij de gemeente?

- Direct dienstverband
- Detachering
- ZZP
- Anders, namelijk

Q22. Selecteer de werktitel die het meest van toepassing is.

- Directeur
- Afdelingshoofd
- Teammanager
- Programmamanager
- Senior
- (Junior) medewerker
- Anders, namelijk

Q23. In welk vakgebied ben je opgeleid? Het is mogelijk om meerdere antwoorden te selecteren.

- ☐ Onderwijs
- ☐ Vormgeving, kunst, talen en geschiedenis
- ☐ Journalistiek, gedrag en maatschappij
- ☐ Recht, administratie, handel en zaken (economie)
- ☐ Wiskunde, natuurwetenschappen
- ☐ Informatica
- ☐ Techniek, industrie, bouwkunde
- ☐ Landbouw, diergeneeskunde en -verzorging
- ☐ Gezondheidszorg en welzijn
- ☐ Dienstverlening
- ☐ Anders, namelijk

Q24. Mijn leeftijd in jaren is ...

- ○ 18 - 24
- ○ 25 - 34
- ○ 35 - 44
- ○ 45 - 54
- ○ 55 - 67
- ○ Anders, namelijk

Q25. Ik ben ...

- ○ Man
- ○ Vrouw
- ○ Non-binair/third gender
- ○ Dat zeg ik liever niet

Block 7

Q26. Wil je ons nog iets meegeven over het onderwerp?

. Ik beoordeel de kwaliteit van deze enquête als ...

- Onvoldoende
- Matig
- Voldoende
- Goed
- Uitstekend

Powered by Qualtrics

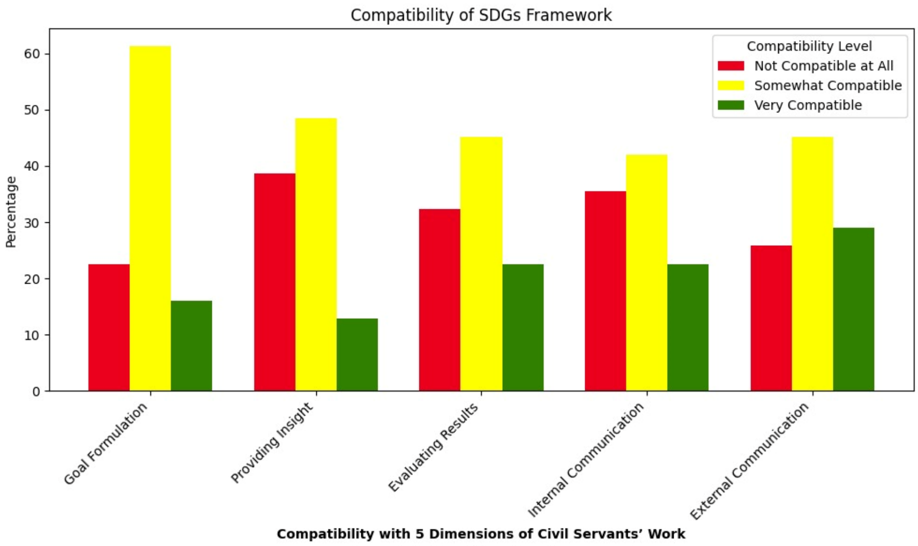
Appendix B. Mann-Whitney U Results

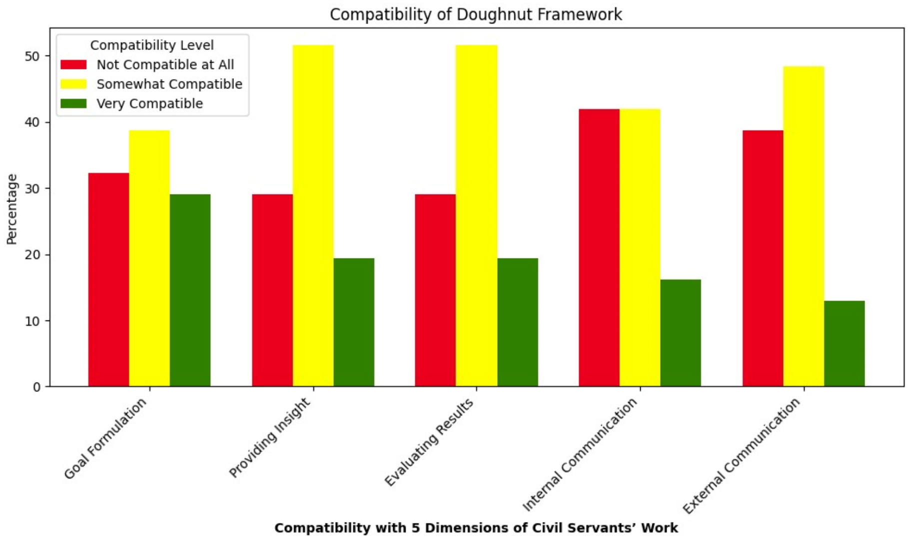
Mann-Whitney U test results

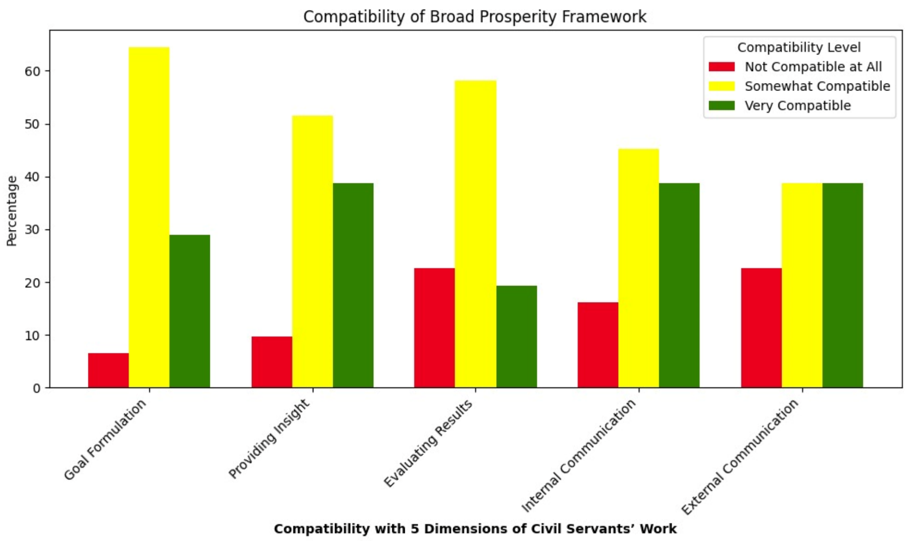
Framework	Variable	U	Groups	
SDGs	Cluster	85.000	Other	Ruimte & Economie
	Policy making	70.000*	No	Yes
	Policy implementation	54.500	No	Yes
	Policy evaluation	99.000	No	Yes
	< City		No	Yes
	City scale	30.500**	No	Yes
	> city, within the country	80.500	No	Yes
	International scale	47.500*	No	Yes
DE	Cluster	89.000	Other	Ruimte & Economie
	Policy making	70.500*	No	Yes
	Policy implementation	52.500	No	Yes
	Policy evaluation	109.000	No	Yes
	< City	80.000	No	Yes
	City scale	18.000***	No	Yes
	> city, within the country	82.000	No	Yes
	International scale	58.500	No	Yes
BP	Cluster	67.000	Other	Ruimte & Economie
	Policy making	98.500	No	Yes
	Policy implementation	62.500	No	Yes
	Policy evaluation	85.000	No	Yes
	< City	99.000	No	Yes
	City scale	69.500	No	Yes
	> City, within the country	56.500**	No	Yes
	International scale	61.000	No	Yes

*Note: N=31 for all three frameworks. *p < .10, **p < .05, ***p < .01.*

Appendix C. Compatibility with five dimensions distributions







Appendix D. Collinearity Diagnostics

E1. SDGs compatibility score on a 1 to 10 point scale = βgoal formulation score + βproviding insights into considerations score + βevaluating results score + βinternal communication score + βexternal communication score

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Correlations Zero-order	Correlations Partial	Correlations Part	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	,815	1,202		,678	,504					
	Q15_1_1	,968	,573	,307	1,689	,104	,531	,320	,241	,616	1,622
	Q15_2_1	,911	,504	,313	1,807	,083	,505	,340	,258	,680	1,471
	Q15_3_1	-,105	,503	-,039	-,208	,837	,266	-,042	-,030	,568	1,760
	Q15_4_1	,776	,538	,299	1,442	,162	,523	,277	,206	,475	2,105
	Q15_5_1	,200	,521	,076	,384	,704	,292	,077	,055	,523	1,912

a. Dependent Variable: Q14_1

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions (Constant)	Q15_1_1	Q15_2_1	Q15_3_1	Q15_4_1	Q15_5_1
1	1	5,608	1,000	,00	,00	,00	,00	,00	,00
	2	,163	5,857	,00	,02	,09	,09	,13	,10
	3	,096	7,648	,00	,00	,32	,25	,10	,11
	4	,058	9,861	,02	,58	,40	,03	,01	,11
	5	,049	10,699	,77	,03	,00	,20	,18	,00
	6	,026	14,623	,20	,37	,19	,43	,58	,68

a. Dependent Variable: Q14_1

E2. DE compatibility score on a 1 to 10 point scale = βgoal formulation score + βproviding insights into considerations score + βevaluating results score + βinternal communication score + βexternal communication score

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Correlations Zero-order	Correlations Partial	Correlations Part	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	1,499	1,099		1,364	,185					
	Q15_1_2	,458	,619	,165	,740	,466	,610	,146	,101	,374	2,674
	Q15_2_2	,427	,559	,136	,763	,453	,479	,151	,104	,589	1,696
	Q15_3_2	1,314	,610	,417	2,154	,041	,652	,396	,294	,496	2,017
	Q15_4_2	,780	,675	,258	1,156	,259	,508	,225	,158	,374	2,674
	Q15_5_2	-,578	,565	-,179	-1,022	,316	,189	-,200	-,139	,609	1,642

a. Dependent Variable: Q14_2

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions (Constant)	Q15_1_2	Q15_2_2	Q15_3_2	Q15_4_2	Q15_5_2
1	1	5,654	1,000	,00	,00	,00	,00	,00	,00
	2	,113	7,072	,00	,03	,14	,05	,07	,31
	3	,105	7,321	,08	,16	,09	,00	,15	,10
	4	,058	9,883	,87	,01	,08	,14	,00	,12
	5	,043	11,498	,04	,03	,55	,79	,02	,01
	6	,027	14,556	,01	,77	,14	,03	,75	,45

a. Dependent Variable: Q14_2

E3. BP compatibility score on a 1 to 10 point scale = βgoal formulation score + βproviding insights into considerations score + βevaluating results score + βinternal communication score + βexternal communication score

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	1,513	1,506		1,004	,325					
	Q15_1_3	,419	,814	,122	,515	,611	,417	,103	,083	,463	2,161
	Q15_2_3	,250	,825	,083	,303	,764	,410	,061	,049	,342	2,920
	Q15_3_3	,940	,720	,321	1,305	,204	,465	,253	,210	,429	2,331
	Q15_4_3	,589	,567	,220	1,040	,308	,398	,204	,168	,583	1,716
	Q15_5_3	,215	,547	,087	,392	,698	,329	,078	,063	,530	1,887

a. Dependent Variable: Q14_3

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions					
				(Constant)	Q15_1_3	Q15_2_3	Q15_3_3	Q15_4_3	Q15_5_3
1	1	5,740	1,000	,00	,00	,00	,00	,00	,00
	2	,133	6,558	,00	,00	,04	,09	,07	,14
	3	,046	11,229	,17	,04	,01	,24	,11	,49
	4	,038	12,284	,23	,04	,00	,09	,78	,14
	5	,030	13,931	,59	,37	,09	,06	,02	,00
	6	,014	20,457	,01	,54	,87	,53	,02	,23

a. Dependent Variable: Q14_3