OECD Working Papers on Public Governance No. 52

# Serving Citizens: Measuring the Performance of Services for a Better User Experience

OECD

OECD Working Papers on Public Governance

# Serving Citizens: Measuring the Performance of Services for a Better User Experience

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcomed, and may be sent to OECD Directorate for Public Governance, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France; e-mail: gov.contact@oecd.org.

This paper was authorised for publication by Elsa Pilichowski, Director, Public Governance Directorate.

# Serving Citizens: Measuring the Performance of Services for a Better User Experience

By Barbara Baredes, OECD Public Governance Directorate

Measuring the performance of services and making effective use of the results are critical for designing and delivering policies to improve people's lives. Improving user satisfaction with public services is an objective in many OECD countries and is one of the indicators in the 2030 Sustainable Development Goal 16 of "Building effective, accountable and inclusive institutions at all levels".

This paper explores the use of satisfaction indicators to monitor citizens' and users' experience with public services. It finds that satisfaction indicators provide an accurate aggregate account of the factors driving service performance. At the same time, it shows that additional measures are needed to monitor the access, responsiveness and quality of public services, as well as to identify concrete areas of improvement.

This paper provides examples of how countries use performance data in decision making (both subjective users' experience and objective service outputs). It also highlights common challenges and good practices to strengthen performance measurement and management.

# Acknowledgements

The author is thankful for the contributions and comments from Monica Brezzi (OECD Public Governance Directorate), as well as the inputs provided by Amira Boland (United States's Office of Management and Budget), Evan Coady (Ireland's Department of Foreign Affairs and Trade), Oscar Jaimes (Mexico's Instituto Nacional de Estadística, Geografía e Informática), Jacqueline Menager (Australia's Department of the Prime Minister and Cabinet), Anne-Sophie Milgram (France's Direction Interministeriélle de la Transformation Publique), Mette Undheim (Norwegian Agency for Public and Financial Mangagement) and Iwona Karwot (European Institute of Public Administration).

Substantive comments were provided by Love Beggrund (Swedish Agency for Public Management), Conor Das-Doyle, Santiago González, Mariana Prats (Governance Indicators and Performance Evaluation Division, OECD Public Governance Directorate), Barbara Ubaldi, Felipe González-Zapata and Benjamin Welby (Open and Innovative Governance Division, OECD Public Governance Directorate), Miyako Ikeda and Corinne Heckmann (OECD Directorate for Education), Valerie Frey (OECD Directorate for Employment, Labour and Social Affairs), and Lara Fleischer (OECD Centre on Well-being, Inclusion, Sustainability and Equal Opportunity).

# Table of contents



## TABLES

# Introduction

People's interaction with public institutions often takes place through the provision of services to individuals such as obtaining an identity card, using services provided by public schools and hospitals, or requesting a social benefit, etc. Citizens' and users' assessments of the performance of the public administration is, therefore, closely related to their experience with and expectations of public services.

Many OECD countries face pressure because of issues such as ageing populations and increased spending due to the COVID-19 crisis. In the near future, governments will need to make substantial investments and transform services to adapt to their population's needs, environmental and financial pressures, as well as societal changes. Digitalisation is also transforming the delivery of services, especially as a result of the COVID-19 pandemic, sometimes without appropriately analysing users' perspectives. Against this backdrop, governments in OECD countries are reconsidering the inputs and tools needed to improve the performance of public services from a citizens' and users' perspective.[1] This raises questions of what are the "right" metrics or methods to measure citizen-centred service performance, what principles and practices public administrations can adopt to create feedback loops between users' experience and improvement of public services, and how to ensure that lessons learned guide decisions on service design and delivery.

An increasing number of countries focus on improving citizens' and users' experience with public services. Users' satisfaction with public services is also one of the agreed indicators to monitor the achievement of the 2030 Sustainable Development Goal 16 of "building effective, accountable and inclusive institutions at all levels". Overall satisfaction with services (e.g. "overall, how satisfied are you with health care services?") has become the go-to indicator when seeking a quick measure of whether services are performing well against user's needs and expectations.

Satisfaction measures have strong links with other relevant measures of citizens' attitudes and behaviour. Satisfaction is linked to trust in public institutions and to the levels of responsiveness and reliability of public institutions (OECD, 2021[1]), (Van de Walle and Bouckaert, 2003[2]).The consequences of dissatisfaction with public services can be multiple, including choosing private providers for services such as health or education (which can reduce the amount of resources available for public services and create segregation in the population by income groups) or deciding to resort to self-administered justice instead of following the legal route (exit behaviour), and working around rules and regulations to enhance own benefit (gaming behaviour) (Peeters, Gofen and Meza, 2020[3]).

Nevertheless, despite the importance of ensuring high levels of satisfaction with services for the good functioning of public institutions, using satisfaction as the measure of whether public services are meeting users' needs and expectations comes with three main challenges. First, the relationship between overall

[1] People-centric public services are those that are accessible to all segments of the population (including the most vulnerable), responsive to citizens' needs and expectations (e.g. to those who are more in need of government support and those who require less assistance and would prefer to use self-service channels) and of high quality (i.e. that deliver on people's expectations and improve their wellbeing) (OECD, 2019[48])).

satisfaction measures and objective indicators of performance (on accessibility, responsiveness, quality and cost effectiveness of services), is not clear-cut, as it is dependent on the objective indicators available, on the policy area of interest, and on a number of external factors. Second, when emitting an overall satisfaction judgement, citizens and users summarise a wide range of aspects of service performance (e.g. the timeliness of the service, the courtesy of the staff, the ease of access, whether they got the result they expected, etc.); hence, its consequent use in decision-making to improve service design and delivery is not straightforward. Third, mistaking the results of satisfaction judgements for the broader objective of delivering on user's needs and expectations (also called "tunnel vision") provides the wrong incentives for decision-makers. This can halt innovation in the public sector, especially if a majority of the citizens and users report being satisfied with services.

Against this backdrop, the contribution of this paper is threefold. First, it discusses the scope and shows the validity of using overall satisfaction measures in monitoring the performance of services. This includes a comprehensive stocktaking of national surveys and measurements of satisfaction with public services.

Second, the paper analyses at a highly aggregated, cross-national level the relationship between objective indicators of performance and satisfaction with services. This paper uses the OECD *Serving Citizens* framework, which compares access, responsiveness and quality of health care, education and justice in OECD countries through a selection of objective indicators. The results show that 81% of the change in satisfaction positively corresponds with a change in 22 objective indicators from *Serving Citizens* (summarised into 5 factors via factor analysis). This indicates that satisfaction can provide an accurate aggregate account of the factors driving service performance (in this case responsiveness, quality and access). In this respect, satisfaction measures are a valid end-target for policies on the improvement of public services, when measured in a way that minimises cognitive biases. At the same time, other performance measures need to be included alongside service satisfaction in order to unpack the different drivers and identify concrete areas of improvement – even when a majority of citizens and users report being satisfied with services.

Third, it discusses common issues for public administrations when utilising users' experience and objective performance evidence to improve the design and delivery of services. This part builds on the experiences and good practices identified during the workshop "Serving citizens: measuring the performance of services for better delivery" organised by the OECD Steering Group of *Government at a Glance* on 4 November 2021.

The paper is organised as follows: Section 1 reviews examples of national policies in OECD countries to improve citizens' and users' experience with a service, many of which include satisfaction as an explicit objective. Section 2 introduces the various ways in which public administrations collect data about satisfaction and provides a literature review of the limitations of such data. Section 3 carries out an empirical analysis to test the relationship between satisfaction with services and objective measures of performance in three services across 13 OECD countries using the OECD Serving Citizens framework. Section 4 discusses how satisfaction data can inform decisions about the improvement of services and some common challenges identified in OECD countries. Section 5 concludes and sets out possible next steps.

# National policies to improve citizens' experience with services

In recent years, many national/federal governments have established policies to improve the design and delivery of services from a citizen-centric perspective. In these policies, citizens are considered co-producers of public services, rather than customers: instead of giving them the choice to change service providers when dissatisfied, as in previous decades, these new approaches give them the opportunity to influence how public services are designed and delivered to ensure that their needs are met. Consequently, to achieve this, governments are seeking to understand how citizens use public services, and to use this information to enhance users' experience with services.

Box 1 shows several examples of policies that strive to improve services for citizens and users. All of these policies link the performance of public administration with outcomes on the population, such as simplified access to services or more personalised services. These policies create feedback loops between citizens' and users' experience and public service design, with the aim of simplifying the interactions between citizens and businesses and their governments. This includes revisiting anything from the language used in communications with citizens and users (e.g. reducing administrative jargon), the design of websites (e.g. ensuring accessibility for the visually impaired) to completely re-designing a service.

### Box 1. National policies to improve the delivery of public services

**Australia: "Delivering for Australians" reform agenda**

In 2019, as a response to an independent review of the Australian Public Service (APS), the Australian government issued the Delivering for Australians reform agenda, which seeks to ensure that the APS is fit for purpose over the coming decades. The agenda focuses on bringing the APS closer to Australians both in decision-making and in the delivery of programmes, projects and public services.

The agenda has six guideposts: "Clear roles and priorities" (ensuring that all civil servants have a clear line of sight of the population they are providing support to), "Better services" (more integration of services, adoption of technology-driven solutions, embedding a customer-focus culture in the APS), "Getting delivery right" (adopting performance targets, tracking the delivery of government priorities and embedding evaluation in the everyday practice of the APS), "Connecting the APS to all Australians" (more engagement with external stakeholders, use of evidence in decision-making and increasing diversity in the APS), "Adapting to change" (reducing bureaucracy, continuing the digital transformation of government, enhancing internal services and human resources development) and "Reinforcing integrity" (strengthening pro-integrity culture including through integrity guidance, processes and tools).

**United States: President's Management Agenda**

In the United States, upon taking office, each new administration establishes a President's Management Agenda (PMA). The PMA is a plan for implementing key management reform priorities and initiatives to improve the effectiveness and efficiency of government. The Biden-Harris management agenda vision – launched in November 2021 – is "an equitable, effective, and accountable Government that delivers results for all." To bring this vision to fruition, the agenda outlines three priorities: Strengthening and empowering the Federal workforce; Delivering excellent, equitable and secure Federal services and customer experience; and Managing the business of Government to build back better.

The second priority on customer experiences outlines three strategies to achieve this vision: improving the service design, digital products, and customer-experience management of designated High Impact Service Providers; managing Federal service delivery through the lens of life experiences that may cross Federal agencies and even levels of government; and prioritising the development of shared products, services, and standards that enable simple, seamless, and secure customer experiences across High Impact Service Providers.

"High Impact Service Providers", are those entities that have a high number of annual transactions, serve a large percentage of the American public, or have an outsized impact on the lives of the individuals that they serve.

**Ireland: Our Public Service 2020**

In 2017, the Department of Expenditure and Reform of Ireland adopted "Our Public Service 2020" as the framework for development and innovation in Ireland's public service. This policy seeks to realise the objectives set out in the Programme for Partnership Government to "empower frontline service providers to make more decisions, ensure more accessible public services, encourage more collaboration between public sector bodies and reward public service innovation and change".

The programme focuses on six high-level outcomes for the public service, namely: Increased customer satisfaction, Increased public trust, Greater use of digital to do business with the public service, Better government effectiveness, Quality of certain public services, and Greater employee engagement. To deliver on them, the programme outlines three key pillars: Delivering for Our Public (this includes actions such as accelerating the digitalisation of services, Improving services for our customers, Making

services more accessible to all, Improving communications and engagement with the public and drive efficiency and effectiveness), Innovating for Our Future (promoting a culture of innovation in the public service, optimising the use of data, building strategic planning capacity, among others) and Developing Our People and Organisations (promote equality, diversity and inclusion, continuous and responsive professional development, among others).

**France: Services Publics+**

In January 2021, the French government adopted the programme "Services Publics+", which aims to adopt a continuous improvement approach to the design and delivery of public services, strengthening the links between users, civil servants and the different administrations. The programme is based on three key themes: closer, more effective and simpler public services. These themes come with three concrete engagements each (e.g. right to make mistakes, more personalised services, respecting the formal deadlines to provide responses to citizens' requests, eco-responsibility) that should be implemented by each public service.

The continuous improvement cycle to implement these commitments includes enhancing the opportunities for users to provide feedback, analysing this feedback (together with users, civil servants and local representatives), displaying and communicating action plans to improve services that emanate from these analyses, executing them, and disseminating the results.

Source: (Australian Government, 2019[4]), (United States' President's Management Council, 2021[5]), (Executive Office of Management and Budget, 2021[6]), (Ireland's Department of Public Expenditure and Reform, 2017[7]), (Direction interministérielle de la Transformation publique, 2021[8]),

These policies do not seek to address how individual services should reach their key goals (e.g. the degree to which health care prevents or treats health problems, or whether social services reduce inequalities), which are usually the objective of sectoral policies and of sectoral monitoring systems. Instead, they address the design and delivery services from a user's perspective. The notion of user design emerged from web design, where there is a clear separation between the "front end" (also known as the "user interface", which are the elements that the user interacts with) and the "back end" (all the processes that run in the background, which users do not see, even though they are the core of the service). Following this distinction, these national policies focus on the aspects that are visible and understandable for citizens and users, and that significantly impact how they perceive the service – even if they also require changes in the back end (such as regulations, processes, division of labour between agencies). This includes issues such as how easy it is to find information about the service, how difficult it is to meet the requirements to use the service, etc.

Naturally, digitalisation is one of the key ways in which governments are seeking to improve public services from a user- and citizen-centric perspective. Many countries have incorporated service design and delivery as a key component of their digital strategies. The OECD Digital Government Policy Framework states that competent digital governments are those that are able to use digital technologies and leverage data about citizens' and users' experience to transform and improve services (OECD, 2020[9]). In this sense, governments are digitally mature when they are able to understand and meet the needs of users in the design and delivery of services (*user-driven*) and anticipate users' needs and provide a seamless and integrated experience across services and channels (*proactiveness*). Nevertheless, the pilot OECD Digital Government Index shows that these two dimensions (*user-driven* and *proactiveness)* are two of the lowest across the six dimensions for OECD countries (OECD, 2020[10]) .

In the quest to harmonise citizens' and users' experience with services, governments face two main issues. First, the public sector provides a wide range of services. As a result, it is challenging to ensure performance is consistent among all services provided, especially given the fact that many services are provided at the local level and are beyond the control of the central/federal government. For this reason,

some policies make an explicit choice to focus on a subset of them, in order to monitor their performance on a regular basis and ensure that they provide a seamless and integrated experience to citizens and users. Some policies provide criteria for prioritising certain services. For example, the United States focuses on "High Impact Service Providers" (see definition in Box 1). Focusing first on services that are widely used by the public can ensure that efforts are focused on those services that can "lift all boats". Stated differently, by embedding good practices in the largest services, a government can gain "spill-over effects" into smaller ones. For example, if a large agency implements a Voice of Customer tool for their High Impact Service Provider, smaller offices can gain efficiencies from this investment. Grouping similar services facilitates the sharing of good practices. Indeed, what works for unpopular services such as taxpaying activities or workplace safety audits will likely be different than what drives satisfaction for other types of services (e.g. farm loan applications or national parks).

Second, services are diverse, as they cover a wide range of individual and social needs and are organised in distinct ways. Drawing lessons from one service to inform decision-making in another one requires identifying points of comparison between them. There are certain services that are more frequently used by certain social groups than by others (e.g. universities are more frequently used by upper classes), and some services are compulsory (e.g. taxes), while users have a choice for others (Van de Walle, 2018[11]). In this regard, there is a tension between the objective of harmonising services and providing users a seamless experience, while catering to the specific needs of the groups that come into contact with each individual service, and delivering on the larger societal goals that these services pursue. In addition, many key services, such as sanitation and refuse collection, are delivered at the local level and are outside of the control of the national or federal government, and the central/federal government can only have a limited impact on them.

Unsurprisingly, making services comparable in order to benchmark their performance and to identify good practices that can be reapplied (with the objective of providing a seamless experience to citizens and users) is one of the main challenges for practitioners[2]. The comparability of services could be enhanced by classifying them by the type of need they address and/or how users come into contact with the service. Box 2 elaborates on the classifications of services used by Mexico and the United States. These classifications make it possible to benchmark and share best practices between services that are included in one category.

[2] Results of the OECD workshop "Serving citizens: measuring the performance of services for better delivery" on 4 November 2021.

### Box 2. Examples of classification of services

**United States: OMB classification of services**

In the United States, the Office of Management and Budget (OMB) published circular A-11, Section 280, which provides the following classification of services:

- Administrative: Requesting or renewing items that do not require an extensive eligibility determination or multi-stage review processes such as getting a license, passport, or social security card.
- Benefits: Applying for or progressing through more complex government processes to determine eligibility and degree of benefit such as immigration, Medicare, Veterans' Health services, or a small business loan.
- Compliance: Completing required actions such as filing taxes, submitting information for or engaging with an auditor, environmental reporting, or completing a survey mandated by law.
- Recreation: Utilising public spaces such as national parks and historical sites, or visiting museums.
- Informational: Providing authoritative knowledge-based resources to the public such as designing labels, releasing warnings, requiring disclosures, or providing health recommendations.
- Data and Research: Conducting or funding research, maintaining and preserving artifacts, collecting, analysing, reporting, and sharing data.
- Regulatory: Providing clear guidance to support commerce, transportation, employment rules, workplace safety, and public safety (e.g. ensuring clean water, safe medicines), which facilitates the reporting of grievances (e.g. consumer protection).

**Mexico: INEGI classification of services**

INEGI's National Survey of Government Quality and Impact (ENCIG) classifies public services as follows:

- Basic public services, such as drinking water, sewers, refuse collection, policing, street lighting, parks and gardens, roads and highways (without toll), streets and avenues.
- On-demand public services, such as compulsory education, tertiary education, health care, electricity, public transportation (buses, metro or light rail), highways and toll roads.
- Formalities, payments and contact with authorities, within these:
  - Payments: property tax, vehicle taxes, payment of electricity and drinking water service.
  - High-frequency formalities: vehicle-related services, taxes, public education and medical appointments.
  - Low-frequency formalities: civil registry, local services (e.g. with municipal government), construction, loans, passports, courts, legal conflicts, starting a business, etc.
  - Request for services: emergency care, municipal utilities, emergency calls to the police.

Source: (Executive Office of Management and Budget, 2021[6]),

Other countries focus on specific "life events" (which are stages of an individual's life when they need to interact with the public administration, such as registering a newborn baby, or getting married) and/or "user journeys" (which refers to the path that a user follows to achieve a goal while interacting with public

services, such as settling in a new home), that usually require the intervention of more than one public agency. User journeys and life events usually sit at the intersection of several relevant public agencies given the division of labour in the public administration. For example, hiring a foreign employee requires obtaining permission from the employment agency, applying for a visa for the employee from the immigration office, etc. Naturally, no individual agency would be responsible for managing this entire journey. Hence, a company trying to hire a foreign employee may find itself requesting a document from an agency, submitting it to a second agency, and so on. More and more governments are realising that this creates unnecessary burden on users, and focusing on providing more integrated services to them.

At the international level, there is currently a certain level of consensus around the main life events of a person, such as the birth of a child, getting married, the death of a loved one, among others. For example, Estonia's "Information Society Development Plan 2020" focused on seven life events that should be handled proactively[3] by the public administration, such as the birth of a child (Republic of Slovenia, 2016[12]). Instead, user journeys are more diffuse and identifying them requires understanding the reasons why an individual or business comes into contact with public services (e.g. they would like to drive a car, they wish to hire a foreign person, they are buying a new house), and reverse engineering the process that the person/company needs to go through to achieve their goal. In these cases, instead of sharing good practices between services (e.g. health care and education), the comparison and sharing of good practices is between life-events/journeys, since one agency (e.g. population registry) can be involved in more than one life event/journey.

In order to ensure all services or journeys are performed to the same high standards and share good practices, a central agency (e.g. the Prime Minister and Cabinet Office in Australia, the *Direction Interministeriélle de la Transformation Publique* in France) is usually assigned with the responsibility for the implementation and monitoring of the central/federal policy for the improvement of services (e.g. Our Public Service 2020 in Ireland, see Box 1). This entails assessing the performance of the different agencies/Ministries that deliver public services (e.g. health care, tax collection, public transportation, etc.) and working with them to improve citizens' and users' experience (e.g. timeliness, courtesy, etc.). These offices are sometimes close to the centre of government (i.e. the structure that serves the prime minister/president and the cabinet), therefore they have a mandate to co-ordinate policies across the central/federal government. In other cases, given the strategic role that digital tools and data play to improve service delivery, this responsibility falls on digital government units. Moreover, digital transformation strategies take the overarching approach of improving services following a user experience approach. This is the case of Brazil, Denmark, UK, Greece, Norway and Portugal.

As part of this mandate, these central agencies are required to provide advice and support to the agencies that deliver public services at the central/federal level on how to gather feedback from users and how to improve their services following citizens' and users' feedback. They must also provide tools that can be reapplied by several services (this can be anything from modules to gather feedback that can be reused in agencies' websites to templates to assess the agencies' capacity to deliver services). The key idea is to support agencies in embedding citizens' and users' feedback in their decision-making processes to ensure that they deliver on the aspects that matter the most to citizens and users[4]. This may involve aspects such as how to develop user-friendly websites, how to use plain language when communicating with users, how to integrate processes with other agencies, among others. This is regardless of the fact that individual agencies may be more adept at understanding what policies and services are needed to achieve their

[3] Proactive services are those that require little to no action by the user (OPSI and MBRCGI, 2020[51]). For example, in the case of the birth of a child, upon registration of the newborn child, parents are prompted to receive child allowances, parental leave, etc. instead of having to request for these services separately.

[4] This also requires incentivising the various agencies responsible for designing and delivering services to co-operate. Ideally, beyond monitoring the performance of agencies to identify good and bad performers, governments could adopt a true "government as a platform" model, see (OECD, 2020[37]).

own goals (for example, the ministry of health may be in a better position to understand how to enhance immunisation and monitor progress).

While there are multiple ways of gathering feedback from citizens and users about their experience with public services, and monitoring how well they are delivering on user's needs and expectations, user and citizen surveys are one of the most common. These surveys have the advantage of gathering feedback that is representative (as they are usually based on representative samples) of citizens and users. Moreover, the feedback is standardised and can thus be quantified and monitored over time. This allows governments to benchmark across services and to quickly assess whether the objective of delivering citizen-centric services is being achieved. In particular, overall satisfaction (e.g. "Overall, how satisfied are you with health care services?") is the go-to indicator when seeking to summarise citizens' and users' experience with public services in a single figure[5]. This also facilitates quick comparisons even when services differ regarding the need that they address (e.g. "x% of citizens are satisfied with health care in comparison to y% with police services, this means that health care services are performing better/worse than police services").

However, there are issues in using an overall satisfaction indicator to compare the performance of services at the central/federal level. First, a single overall satisfaction judgement does not indicate what standards an individual is using to assess their experience or whether they are comparable across services. For example, to assess their satisfaction with health care, a person may focus on how clean the health care facilities were, while for education the same individual may emphasise their child's grades when evaluating their satisfaction with education. Second, a persons' reported satisfaction may be more revealing of the subjectivity of the person who emits the opinion (e.g. their expectations or their prejudices, which relate to the person's socioeconomic background and experiences) than of the performance of the service itself. Similarly, overall satisfaction indicators do not provide granular or actionable insights that can drive improvements in service delivery.

In the quest for simplicity and comparability, some policies go as far as to confound the objective of improving citizens' and users' experience with the objective of improving overall satisfaction ratings. This narrow focus on satisfaction could lead to oversimplification (e.g. not looking beyond averages to understand whether there are particularly dissatisfied groups) and thus obscure the objective of responding to citizens' needs, if this indicator is not measured and interpreted properly. Using citizen satisfaction results as a single proxy for measuring whether services are meeting citizens' and users' needs and expectations can create ill-advised incentives. For example, if a majority of citizens report being satisfied with the current functioning of services (or specific attributes, such as image and trust in existing services) governments may become comfortable with the status quo and not pursue actions to improve service design and delivery (e.g. moving towards proactive services).

The following section provides a literature review of satisfaction with services, describing what it measures, the potential biases that satisfaction measurements carry and how they can be avoided.

[5] For example, when reporting on the SDG 16.6.2 "Proportion of the population satisfied with their last experience of public services", some countries report on a single metric for all services combined. This is the case for Germany (https://sdg-indikatoren.de/en/16-6-2/).

# Measuring satisfaction with public services

Satisfaction is a subjective judgement of service performance, affected by a person's experience of the service process (e.g. ease of access) and the outcome of such service (e.g. Did the person get what they were entitled to?). As a subjective measure, satisfaction can only be reported by the person who consumed the service and it reflects how the person interpreted their experience.

From the 1970s, with the advent of the New Public Management (NPM) paradigm, governments sought to increase the efficiency and effectiveness of public spending, particularly focusing on measuring the relationship between outputs (e.g. number of treated patients) per input (e.g. spending on health care, number of health care personnel). NPM shifted the centre of the public administration's attention from its own inner processes to citizens and users and their expectations.

These transformations included the introduction of citizen-surveys. As services liberalised in the 1980s and 1990s, complaint mechanisms were often set up. In addition citizens and customers charters were implemented to set standards that public organisations had to reach and maintain (Van de Walle, 2018[11]).

More recently, satisfaction with public services has become a key indicator of the quality of governance in the international agenda. The Praia Handbook of Governance Statistics provides guidance to national statistical offices on the development of governance indicators, following international standards. To this end, the handbook breaks down the concept of governance into eight dimensions. One dimension is responsiveness, which is defined as "the degree to which public institutions listen to what people want and act on that" (Praia City Group, 2020[13]). This is measured by citizens' satisfaction with public services and political efficacy. System responsiveness is a key feature of democratic systems, given that it reflects the capacity of citizens to bring about change. This capacity also contributes to ensure citizens' commitment and participation in democracy.

As a result of these efforts to deliver for citizens, many mechanisms for collecting feedback and satisfaction data from users have become omnipresent. Most countries now have a complex system of surveys to gather data on satisfaction, as is the case in Norway (see Box 3), where there are surveys by individual service providers as well as central surveys. Other mechanisms to collect data on citizens' and users' experience and perception of public services include feedback mechanisms. For example, Bogota's Veeduria Distrital developed a dashboard of citizen feedback[6]. More recently, user research has emerged to understand the impact of design on an audience (Goodman, Kuniavsky and Moed, 2012[14]) drawing on qualitative research techniques, such as focus groups and non-participant observation, and from design techniques, such as design thinking. In the United Kingdom, virtually every government department has

[6] http://tablerocontrolciudadano.veeduriadistrital.gov.co:3838/TCC/

operational researchers, whose role is to analyse service delivery and to seek improvements following user research techniques[7].

### Box 3. Surveys on citizens' experience with services in Norway

In Norway, multiple surveys on satisfaction with services exist. Since 2010, the Norwegian Agency for Public and Financial Management (DFØ) has conducted the Citizen Survey, which aims to collect data on citizens' experience and satisfaction with services. Respondents that had used a particular service (e.g. hospitals), were prompted to answer a specific "User Survey" that enquired about their experience with the service. This included aspects such as quality, accessibility, benefits, information and communication, consumer orientation, competence, trust, and overall satisfaction.

Given the low response rate for the majority of services, in 2017 the DFØ decided to focus on a subset of them. By 2019, only the modules on hospitals, tax administration, police, and student loans remained. Instead, since the start of the Citizen Survey, individual providers of services had established their own user surveys that measure customer satisfaction, and focus on attributes that are specific to their own service. These surveys have a higher response rate than the User Survey, which provides more solid analyses than the DFØ could provide. For this reason, the DFØ decided to discontinue these detailed modules; instead, the Citizens Survey provides general data on satisfaction with each service and compares these ratings between those who have experience with a service and those who do not.

The results of the OECD Digital Government Index show that countries have adopted citizen satisfaction measurement frameworks to enable users to communicate about their expectations and provide feedback on the efficiency and responsiveness of services to their needs. In 2019, following the Digital Government Index, 51% of countries were assessing user satisfaction through indicators, 49% had guidelines for measuring user satisfaction with digital services, and 12% had made them mandatory across levels of government (OECD, 2020[10]).

Table 1 shows OECD countries' surveys on citizens' and users' experience with and/or perceptions of public services. The majority of these are conducted every two years or more frequently. For example, Australia's Citizen Experience Survey is collected quarterly. Some countries only occasionally include a module on public services in surveys that focus on other topics, such as the social survey from Israel's Central Bureau of Statistics that measures citizen wellbeing. These data are used to understand whether services are meeting citizens' and users' needs and expectations. They are popular because it is easy to interpret them and identify directionality, which allows the government to draw quick conclusions and inform high-level discussions on how to improve public services.

The responsibilities for data collection vary between surveys. Some of these surveys are managed by national statistical offices (e.g. Destatis in Germany), while others are under the responsibility of think-tanks (e.g. Institute for citizen-centred service in Canada). These agencies do not have decision-making power over the delivery of services, hence their role is to supply evidence on their performance without following up on their use. In particular, under the UN Fundamental Principles of National Official Statistics, national statistical offices are expected to be insulated from the influence of the government to be able to provide impartial statistics (OECD, 2015[15]). Other surveys are directly managed by an institution mandated to improve the performance of services (such as the Modernisation Secretariat within Chile's Ministry of Finance). In some cases, line ministries and agencies conduct their own surveys following guidelines provided by the institution that is responsible for overseeing the improvement of services (e.g. the OMB in the United States).

[7] For more information visit: https://www.gov.uk/service-manual/user-research

In terms of services covered, some of the surveys displayed in Table 1 focus on services by several levels of government, such as the Quality of Government survey from Mexico that covers anything from services provided at the local level (e.g. street lighting) to those provided at the federal level (e.g. birth registration). Other surveys focus only on services provided at a specific level of government. For example, Colombia's Planning Department has a survey that focuses only on services provided by the national level, while Italy's municipalities have surveys focusing on services delivered at the local level.

These surveys are usually in accordance with the public policies that underpin them. An example is how they address the digital transformation of the public sector. Some of the surveys displayed in Table 1 focus only on digital services, meaning those that are provided through digital channels. This is the case of Switzerland's National eGovernment Study that supports the monitoring of the eGovernment Strategy of Switzerland[8]. It seeks to monitor the penetration of digital government in society and businesses, and assess users' level of trust and satisfaction with online services. In other cases, digital channels are considered alongside other ways of accessing services. This distinction reflects the fact that some countries have adopted "digital first approaches", while others focus on how to deliver a similar or uniform experience through various channels, supported by digital technologies ("omni channel approach"). This is the case in Portugal where, since 1999, physical "one stop shops" called Locais de atendimento[9] have been set up. At the shops, citizens can access a wide range of public and private services in a single place. The shops use the same digital infrastructure (e.g. databases and data sharing arrangements) as the digital central service delivery platform ("digital one stop shop").

Despite the importance of understanding citizens' and users' perspective of the performance of services, subjective indicators of performance, such as satisfaction, are 'close to the unit of analysis' – in other words, people's views may be biased. This is in contrast with more objective indicators of performance that are detached from the unit of analysis, and require a clear definition of the dimension of performance that is being measured (Andrews, Boyne and Walker, 2006[16]).

According to the expectancy-disconfirmation theory, people rate their satisfaction with a service by contrasting their experience with prior expectations (Van Ryzin, 2004[17]). Expectations can be predictions, anticipations or beliefs about the quality of a service, or they can be normative assertions about what the service should be (Poister and Thomas, 2011[18]). Expectations are highly subjective and may vary between individuals, even when they receive a similar service. For example, respondents who interact more frequently with the government (and who possibly have higher awareness of their own rights and of their government's obligations) have higher expectations of what constitutes good quality compared to the rest of the population (UNDP Oslo Governance Centre, 2020[19]).

External cognitive biases can also impact how individuals rate public services in surveys, especially those that that they have not used. These biases include "anti-public sector" bias, which is the tendency to rate public services worse than private ones even if they perform similarly on objective metrics (Marvel, 2015[20]). Along the same lines, political opinions (e.g. support of or opposition to the political leadership) can influence people's ratings of public services (see (Ipsos-MORI, 2006[21]) on satisfaction with the management of the UK National Health Care system). In addition, halo effects (i.e. an overall impression overshadows more specific traits) can lead citizens to not differentiate between the performance of one service from another (Van de Walle, 2018[11]). Finally, negativity bias (i.e. given events or reports of a similar intensity, those of a negative nature have a larger impact on a person's judgement or feelings than neutral or positive ones) is a pervasive trend, where one bad experience outweighs a positive one in citizens' ratings (e.g. (Li, Ren and Luo, 2016[22]); (Olsen, 2015[23]).

[8] https://www.egovernment.ch/en/umsetzung/e-government-strategie/

[9] https://eportugal.gov.pt/locais-de-atendimento-de-servicos-publicos

However, the literature shows that direct experience with services can help citizens and users to make a more informed judgement and to take a different set of dimensions (more related to the service itself) into account when rating their satisfaction (Van de Walle, 2018[11]). Other studies show that information about the performance of services (e.g. scorecards by the local government) affects citizens' expectations (James, 2011[24]).

Questionnaires can be designed to steer citizens' and users' assessments away from cognitive biases, and nudge them to focus on the performance of services. Two ways of reducing cognitive biases include explicitly asking individuals to recall their latest experience with public services, also known as anchoring (Andersen and Hjortskov, 2015[25]). Then individuals are asked about the characteristics of the service before asking them to share their overall satisfaction level, which controls for question-order bias (Grimmelikhuijsen and Porumbescu, 2017[26]). For example, the methodology of the SDG 16.6.2 "Proportion of population satisfied with their last experience of public services" takes this approach by asking respondents to reflect on their experience with services before indicating their overall satisfaction (UNDP Oslo Governance Centre, 2020[19]).

Still, the biases in measuring satisfaction have led some scholars to criticise the use of satisfaction as the main source of information to make decisions about budget allocations or changes in service design and delivery (e.g. (Howard, 2010[27]). Further, experts in a field may have their own idea about how to best achieve the individual and social goals of a specific service in response to evidence that may be available. However, normal people may have their own beliefs or preferences that contradict expert views (e.g. some people may believe that some content should not be taught at school, and may become dissatisfied if their children study such content).

Instead, the OECD's *Government at a Glance* publication suggests to complement satisfaction ratings with data collected by other sources, such as administrative data, to gather a complete picture of service performance. For example, a review of studies comparing administrative data and surveys showed that the former reflect the views of regulators, while surveys reflect the views of citizens and users (Andrews, Boyne and Walker, 2011[28]).

## Table 1. Countries' surveys on satisfaction with services

| Country | Institution | Survey | Rationale | Population | Services covered | Attributes | Methodology | Frequency | Website |
|---|---|---|---|---|---|---|---|---|---|
| Australia | Department of the Prime Minister and Cabinet | Citizen Experience Survey | The Survey explores citizens' experiences of Australian public services during life events, to improve services at a whole-of-government level and increase peoples' trust and satisfaction in services. | Individuals | Public (also known as federal, national, Commonwealth level) services, including passport applications, citizenship as well as elderly care, employment, tax and welfare payment services, related to approximately 50 life events. | Overall trust and satisfaction in public services, and satisfaction drivers (accurate, transparent, accountable and respectful); experience drivers (effort and time is reasonable, clear processes, easy access); and trust drivers (reliable, responsive, fair, open and honest, integrity) | Nationally-representative monthly survey of 1,000 respondents aged 18 and over. Standardised questionnaire using a private online panel. Results reported quarterly (~3,000 responses). | Quarterly, from Feb 2019-June 2021, monthly since July 2021 | https://pmc.gov.au/public-data/citizen-experience-survey |
| Canada | Institute for citizen-centred service | Citizens First | Understand drivers of satisfaction by delivery channel, measure citizens' expectations and explore the relationship between service delivery and trust in government | Individuals | A wide variety of services delivered at the national, provincial and municipal level (e.g. sewage, recycling, fire department, libraries, museums, courts, roads, tax payment, obtaining licences) as well as the variety of channels (online, in-person, telephone service) | Service design, delivery timeliness, staff interaction, channel functionality, issue resolution | Standardised questionnaire administered to a randomly-recruited internet panel, as well as mail and mail-to-online (respondents received the questionnaire by mail, but were allowed to provide online answers). Two samples, one of 800 respondents representative at the national level, another representative at the jurisdictional level of 400 respondents per subscribed jurisdiction. | Every two years | https://citizenfirst.ca/research-and-publications/citizens-first |
| Chile | Secretaría de Modernización - | Programa de Modernización | Understand drivers of satisfaction by type of | Users (individual or | A wide range of services that count | Ease (e.g. timeliness, number of procedures) | Standardised questionnaires | Every two years | https://www.satisfaccion.go |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Ministerio de Hacienda | del Sector Público | user, type of service and channel. | institutional) | 80% of the demand of services in the country without considering public health and education sectors. | Likeability (treatment, clarity of requirements), Responsiveness (solved the problem), Image (trust in institution) | administered by phone and online, representative sample of the user base of each service. | | b.cl/ |
| Colombia | Departamento nacional de planeación | Encuesta de Percepción Ciudadana | Understand citizen's views of the quality and accessibility of government services, as well as their interests, expectations, and needs regarding public entities. | Individuals | Administrative services (paperwork) | satisfaction drivers quality, waiting times, cost, courtesy, clear communication, agility and speed, channels; barriers | Sample of 9,926 interviews in 10 cities, in-person household survey. Random stratified sample. Population aged 18 years old and older. | Yearly | https://www.dnp.gov.co/programa-nacional-del-servicio-al-ciudadano/Paginas/Encuesta-de-Percepci%C3%B3n-Ciudadana-.aspx |
| Denmark | Statistics Denmark and Digitaliseringsstyrelsen | It-anvendelse i befolkningen (IT use in the population) | Understand the use of digital technologies in the Danish population (by groups) | Individuals | Digital services (both private and public) | Confidence in government digital services, use of "welfare technology" (e.g. health), digital communications with public authorities, use of digital public services. | Sample size 5,800 respondents. Population aged 15-89 years. Telephone interviews. | | https://www.dst.dk/da/Statistik/nyheder-analyser-publ/Publikationer/VisPub?cid=29450 |
| France | Institut Delouvrier and Délégation Interministérielle de la Transformation Publique (DITP), placed under the authority of the Ministre de la Transformation et de la Fonction publiques | Baromètre de l'Institut Paul Delouvrier | Understand people's expectations of public services, perform a comparative evaluation of the missions of the various public services, analyse the relevance of the different channels | Individuals | Police, health care, education, environment, social security, justice, tax authorities, housing, employment | Timeliness, promptness, responsiveness, treatment, ease of administrative procedures, affordability, confidentiality, among others, depending on the service. | Standardised questionnaire administered online to a nationally representative sample of 2,500 individuals. | Yearly | https://www.modernisation.gouv.fr/etudes-et-referentiels/la-satisfaction-des-usagers-des-services-publics-sameliore-en-2020-les-enseignements-du-barometre-2020-de- |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | linstitut-paul-delouvrier |
| | DITP | Baromètre de la complexité | Evaluate the level of complexity experienced by users, understand how citizens interact with the public administration through these life events and measure pain points for users | Individuals | 25 life events | General satisfaction, perception of complexity, channels, accessibility of information, receptivity and treatment of the demand. | Standardised questionnaire administered over the phone. Quota sample of 7,700 individuals aged 15 or older. | Every two years | https://www.modernisation.gouv.fr/publications/barometre-2020-de-la-complexite-administrative |
| Germany | The Federal Statistical Office, commissioned by the Federal Government | Life situations survey | Contact with the public administration is primarily assessed from the perspective of individual life situations and the resulting requests and duties. Regular life situation surveys on the satisfaction of citizens and companies with public authorities and their service. | Individuals and enterprises | 21 life situations (life events) for individuals.<br>10 life situations for entreprises. | Trust, Competence, Courtesy, Responsiveness, Availability and accessibility of the information, Simplicity, Websites, Equal treatment, among others. | Standardised questionnaire administered over the phone. Sample of 5,500-6,000 individuals aged 16 or older who had contact with the public administration in the framework of the 21 life situations. Sample of 2,600 interviews for companies represented by a manager, which experienced at least one of 10. | Every two years | https://www.amtlich-einfach.de/DE/Ergebnisse/Buerger2021/Ueberblick_node.html |
| Germany, Austria, Switzerland | Initiative D21 | E-Government Monitor | Analyse the situation of the digital administration in Germany, Austria and Switzerland | Individuals | E-government services | Awareness, utilization, satisfaction and usability, barriers and attitudes | Online panel. Representative quota sample (gender, age, education) of 1,005 individuals in Germany, 1,008 in Austria, and 1,002 in Switzerland | Yearly | https://www.egovernment.ch/en/dokumentation/controll/ |
| Ireland | Department of Public Expenditure and Reform | Survey of Civil Service Customer Satisfaction Survey Results | Understand drivers for satisfaction (and dissatisfaction) with the civil service and by channel, assessing attitudes | Individuals | Government departments/offices (i.e. ministries) | Speed, efficiency, outcome, non-response | Standardised questionnaire conducted face-to-face in 170 sampling points in Ireland, nationally representative sample | Every two years, since 1997 | https://www.gov.ie/en/policy-information/a38d80-civil-service- |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | and perceptions of the civil service. | | | | of 2,027 individuals (quotas). Respondents aged 18 and above. | | general-public-customers/ |
| | | Civil Service Business Customer Survey | Determine levels of satisfaction with, and perceptions of, the services provided by the Civil Service Departments and major Offices.by channel | Entreprises | Government departments/offices (i.e. ministries) | Speed, efficiency, outcome, non-response | Standardised questionnaire conducted over the phone. Stratified random sample of 510 entreprises, with quotas by number of employees and location. | Every two years | https://www.gov.ie/en/press-release/36a072-results-of-2018-survey-of-civil-service-business-customers/ |
| Israel | Central Bureau of Statistics with Budget Division of Ministry of Finance | Social Survey | Understand the well-being of Israeli population | Individuals | Wide range of services (e.g. parks, health care) | | Sample of 7,078 people aged 20 and above. | Yearly | https://www.cbs.gov.il/he/publications/Pages/2017/ |
| Italy | Municipalities (e.g. Florence, Rome) | Citizen Satisfaction Surveys | Understand the needs of citizens to improve services (http://www.funzionepubblica.gov.it/articolo/dipartimento/24-03-2004/direttiva-customer-satisfaction) | Individuals | Municipal services (e.g. library, cemetery, social services) | Staff (competence, courtesy), timeliness, clarity of information | | | http://www.magalab4.com/limesurvey/index.php/884668?lang=it |
| Japan | Cabinet Office | Public Opinion Poll | Understand the public's awareness of public services | Individuals | Services in general | Satisfaction with services in general and in particular. Drivers of dissatisfaction (e.g. complicated to use, not commensurate with taxes paid) | Sample of 1,727 respondents, 2-stage stratified random sample. People aged 20 and over. In-person interviews. | Yearly (although the one on public services was in 2015). | https://survey.gov-online.go.jp/index.html |
| Latvia | SKDS for the State Chancellery | Public Administration Client Satisfaction Survey | Measuring level of satisfaction of the Latvian population by evaluating the work of public administration | Individuals | Services in general | Improvement, Simplicity, Channels, Use and rating of public services, Evaluation of co-operation within the administration | Sample of 1,000 respondents, random stratified sample. People between the ages of 18 and 75. In-person household | Yearly | http://petijumi.mk.gov.lv/node/3873 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | institutions, the attitude towards users and the quality of the services provided. | | | | interviews. | | |
| Mexico | INEGI | Encuesta Nacional de Calidad e Impacto Gubernamental | Understand the characteristics of public services, estimate the time dedicated to filling out paperwork related to public services, measure perceptions and experiences of corruption, measure trust in government | Households | A wide variety of services classified as basic public services (e.g. street lighting), public services on demand (e.g. education), payments (e.g. electricity bills), high-frequency formalities (e.g. driver's licence), low-frequency formalities (e.g. birth registration), service requests (e.g. emergency requests to police), and authority acts (e.g. arrest and detainment). | Timeliness, staff interaction, quality of public goods and infrastructure, safety, consistency, affordability, among others | Standardised questionnaire administered in person. Sampling of households in cities of more than 100,000 inhabitants, representative at the national level and at the jurisdiction level. A total of 46,000 interviewed households. | Every two years | https://www.inegi.org.mx/programas/encig/2019/ |
| Norway | Agency for Public Management and eGovernment (DIFI) | The Norwegian Citizen Survey (Innbyggerundersøkelsen) | The Norwegian Citizen Survey is both a citizen survey and a customer satisfaction survey. It is one of the largest surveys of public services in Norway. | Individuals | Population section: Provides an overall picture of the inhabitants' view of the municipality of residence, including an assessment of the municipal services and trust in politicians and the administration.<br><br>User section: maps experiences with 22 selected public | Quality, accessibility, benefit, information and communication, consumer orientation, competence, trust, and overall satisfaction | Standardized questionnaire administered by mail to ¼ of the sample and by e-mail to ¾ of the sample. Random sample of 7,100 individuals aged 18 and above. | Every two years | https://www.nsd.no/nsddata/serier/innbyggerundersokelsen_eng.html |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | services in the areas of education and culture, health, care and government agencies. | | | | |
| Spain | Observatorio de Servicios Urbanos (OSUR) | Barómetro de Satisfacción de los Servicios Urbanos | Gauge levels of satisfaction with public services | Individuals | Municipal services (water, transportation, refuse collection, electricity and gas, parks, cleaning) | | Sample of 5,500 respondents at the national level in the 30 most populated urban areas of the country. | Yearly | https://www.osur.org/barometro/#barometro-de-satisfaccion |
| | Dirección General de Gobernanza Pública and Centro de Investigaciones Sociológicas | Calidad de los Servicios Públicos | Identify areas of improvement of public services | Individuals | Public services in general, and in particular (education, health care, social protection, transportation, infrastructure, justice, security) | Simplicity, clarity of information, timeliness, use of digital technologies, courtesy and treatment, facilities, staff professionalism, possibility of participation, online access | In-person household interview, people aged 18 and older. Sample of 3,342 interviews. | Yearly | https://www.mptfp.gob.es/portal/funcionpublica/gobernanza-publica/calidad/informes/percepcion.html |
| Switzerland | DemoSCOPE with eGovernment Switzerland Programme Office | National eGovernment Study | Understand the development of e-government services in Switzerland | Individuals, businesses, public administration | E-government services | Awareness, utilization, barriers, support provided by authorities, satisfaction with services, trust in services, | Online and over-the-phone interviews. Sample: 2,606 individuals aged 18 to 70 (general population), 1,331 businesses located in Switzerland (secondary and tertiary sectors), 27 federal offices and departments, 26 cantons and 1,010 municipalities. | Every two years | https://www.egovernment.ch/en/dokumentation/controll/ |
| United States | Office of Management and Budget | Designated High Impact Service Providers are required to offer post-transaction surveys in line with government- | HISPs designate priority services (those with a high volume of annual transactions, serve a large percentage of the American public, or have an outsized | Dependent on service; primarily individuals or businesses | Each HISP (currently 35 entities) required to designate at least two | Overall: Trust, Satisfaction<br>Drivers: Service Effectiveness / Perception of Value; Ease / Simplicity; Efficiency / Speed; Equity / Transparency; Employee | Varies by HISP and service; volume of transactions, surveys presented, and surveys completed are included in data reporting on performance.gov/cx | Real-time post transaction; results reported quarterly | https://www.performance.gov/cx/ |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | wide guidance | impact on the lives of those that they serve) to map customer journey and instrument feedback points along | | | Interaction / Warmth / Helpfulness / Competence | | | |

# Relationship between satisfaction and objective indicators of the performance of public services

Given the relevance that satisfaction measures have acquired in making decisions on the improvement of public services, understanding the links with other types of measures of performance that governments use to assess the impact of services on citizens' well-being is key. This section carries out an empirical analysis to investigate the relationship between satisfaction with services and objective indicators of service performance, making use of the indicators and data available in the OECD Serving Citizens framework, which features since 2015 in the publication OECD *Government at a Glance*.

The Serving Citizens Framework (Table 2) evaluates three dimensions of public service performance from a people-centric point of view: access[10], responsiveness[11] and quality[12]. These dimensions are comparable across services and the framework seeks to explain satisfaction with services.

[10] Access can be thought of as the opportunity to obtain an appropriate service in case of need (Levesque, Harris and Russell, 2013[43]), and the indicators cover the aspects of affordability, geographic proximity and how easy it is to access information. In recent years, the channels through which services are provided have diversified, including call centres and digital channels on top of physical offices. Hence, beyond the geographic proximity, it has become important to assess whether it is possible to access a service through a user's preferred channel, and to switch between channels without having to start the process over again. Easy access to information also refers to whether the information can be easily found, understood and whether it is feasible to clarify the information in any way (e.g. having FAQs, etc.).

[11] Responsiveness refers to how quickly and well public organisations respond to people's expectations. This implies that public services take into account the needs, preferences, perspective and dignity of individuals who use them, and are provided without unreasonable delay. This includes the aspects of courtesy and treatment, match of services to special needs (i.e. whether service providers adapt the delivery to the different segments of the population, such as for people with disabilities), and timeliness.

[12] Quality is the degree to which services increase the likelihood of desired outcomes and are consistent with current professional knowledge (Kelley and Hurst, 2006[49]); the indicators cover aspects of effective delivery and outcomes, consistency in service delivery and outcomes, and security (safety). Consistency in the delivery refers to degree of provision of the same service and the achievement of the same outcomes across all segments of the population. Security/safety entails that people are not exposed to unnecessary risk let alone harmed or injured in the service delivery process. Whereas the definition of quality stresses the achievement of desired outcomes, the sub-dimension of safety covers the unintended or side effects of public services.

## Table 2. The Serving Citizens Framework

| Access | Responsiveness | Quality |
|---|---|---|
| Affordability | Courtesy and treatment | Effective delivery of services and outcomes |
| Geographic proximity | Match of services to special needs | Consistency in service delivery and outcomes |
| Accessibility of information | Timeliness | Security (safety) |

Since 2015[13], the Serving Citizens Framework has been applied to such diverse areas as health, education and justice through a selection of sector-specific indicators defined by relevant OECD committees, and selected by sector experts to feature in *Government at a Glance* (GAAG). The selection of indicators reflects the most up-to-date knowledge about the performance of services from an expert perspective. The indicators used in *Government at a Glance,* which measure the performance of services following the dimensions of the framework, combine objective indicators (coming mostly from administrative sources) and experience-based indicators (coming from surveys to the population, which are presented alongside satisfaction with services).

The Serving Citizens Framework is consistent with other international efforts to measure the citizen-centricity of services at a macro level (i.e. comparing across countries and years). This includes the Sustainable Development Goals indicator 16.6.2, "Proportion of the population satisfied with their last experience of public services" (UNDP Oslo Governance Centre, 2020[19]) and the European Institute of the Public Administration's (EIPA) report comparing the performance of public services in 35 countries to be published in 2025 (See Box 4).

[13] To measure the Serving Citizens Framework, a subset of sector-specific measures from education, health and justice are selected by experts from the OECD on each subject for each edition of GAAG, according to five criteria. The criteria are: adequacy (i.e. the indicator represents the concept being measured), policy relevance, data availability and coverage, comparability; and data interpretability (i.e. no ambiguity that a higher/ lower value means better/worse performance). The selected indicators are intended to provide an overview of relevant aspects for each sector and represent a small subset of those included in specialised publications (such as OECD Health at a Glance, PISA and Education at a Glance). The choice of indicators differs among sectors (e.g. school enrolment for education and health care coverage for health care are measures of access). The indicators to measure a specific dimension of public services may vary from one edition of GAAG to another, because of differences in data collection frequencies of international data sources (e.g. PISA, Health at a Glance), countries' coverage of internationally comparable sources (e.g. Eurostat and the Commonwealth Fund), or refinement of measures for the underlying concepts. The changes in the selection of indicators also reflect the programme of work of each committee, following new developments in their sectors.

### Box 4. European Institute of the Public Administration's study comparing the performance of public services across 35 countries

The European Institute of Public Administration (EIPA), in co-operation with the Dutch Ministry of the Interior and Kingdom Relations, will conduct the "International Benchmarking Study 2022 - 2025: Public Sector Performance''. The study provides in-depth comparative analysis of public services' performance in 35 countries (27 EU Member States, United Kingdom, Norway, Iceland, Switzerland, Australia, New Zealand, Canada, and United States) across 10 policy areas. The research will be structured as follows: (i) the sub-study in 2022 will focus on Public Administration, Education, Housing and Social Safety; (ii) the sub-study in 2023 will cover Economy, Infrastructure and STI, Social Security, Employment, Income and Wealth, Environmental Protection and Climate change; (iii) the sub-study 2024 will address Health, Sports, Culture & Participation, and a 10th policy area that will be selected in mutual agreement with the Ministry. In 2025, all policy areas will be updated to reflect the developments that may occur in the ensuing years.

The overarching objective of the study is to evaluate the performance of public services, their (cost)-effectiveness and efficiency, the satisfaction level and trust of citizens, enterprises and other societal actors towards their public administrations. Where applicable, the study will analyse mechanisms, regulations, policy-systems, and other relevant aspects or contexts that have an impact on the success or failure of public services. The ultimate aim of the research is to provide policy makers with the opportunity to learn about good practices and the most important factors that guarantee the successful functioning of public services and a high level of achievement of policy objectives.

In the literature on the performance of services, the link between objective measures (for example, measures that come from administrative data, such as enrolment rates in education or health care coverage, which are process and output measures) is a subject of debate. In an analysis of the performance of the police of Los Angeles, Stipak (1979[29]) found weak relationships between objective indicators, such as clearance rates, and citizens' evaluations. Similarly, (Swindell and Kelly, 2002[30])tested the relationships between the objective performance measures of local services (police, parks and recreation, fire and emergency department) in the United States as defined by the International City/County Management Association and found less-than-expected correlations with citizens' satisfaction ratings.

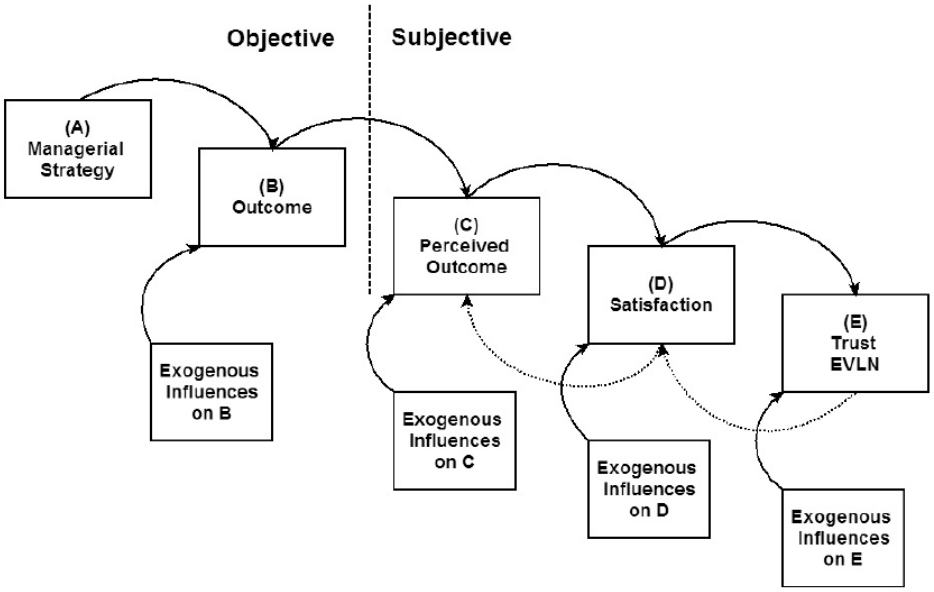
A recent analysis of the OECD's Risk that Matters survey, which focuses on people's perception of social and economic risks and their public policy preferences, shows ambiguous results. A large proportion of perception-based indicators (e.g. quality and affordability of long-term care) showed a high correlation with objective indicators (e.g. long-term care spending) (OECD, 2019[31]). However, in 2021, a significant proportion of respondents in countries like France believed that the government wouldn't provide enough support to see them through financial difficulties – despite the government covering up to 70% of the average wage for those at risk of unemployment due to the COVID-19 lockdowns (Frey, 2021[32])

Parks (1984[33]), using path analysis, was able to identify relationships between objective variables of police performance, such as the existence of patrol units, and citizen experience measures (e.g. whether the person saw a patrol car), and compare them with levels of satisfaction with the police in the Saint Louis metropolitan area in the United States. He identified that a change in objective performance affected citizens' perceptions of police performance. Similarly, Van Ryzin, Immerwahr and Altman (2008[34]) correlated the responses to a citizen survey on the cleanliness of their street with the New York City's street cleanliness scorecard, and found a strong relationship. They found the scorecard to be a stronger predictor of citizens' assessments than trust in government, contextual factors or demographic variables.

Van Ryzin (2007[35]) modelled the relationship between objective and subjective indicators. In his view, managerial strategies affect outcomes (such as the cleanliness of the street, which can also be thought as

outputs), which affect citizen perceptions of these outcomes (or outputs). These perceptions affect satisfaction with services, which in turn affects trust in government. However, for subjective indicators there is also a reverse relationship. All subjective indicators are affected by exogenous factors; hence, there is no perfect correlation between any of the elements of this model.

### Figure 1. Relationship between government performance, citizen satisfaction and trust in government




Note: EVLN refers to exit (e.g. leaving a jurisdiction), voice (complaining), loyalty (trust), and neglect (apathy or alienation), which are the reactions to satisfaction or dissatisfaction.
Source: (Van Ryzin, 2007[35])

Consistent with this model, Norris (2022[36]) analysed the relationship between citizens' trust in government and expert-based measures of good governance from international sources, and variations in the results according to the national context. In societies with freedom of expression and widespread access to alternative sources of information, there is a strong correlation between people's perception of their government and expert-based measures of performance. In contrast, for societies with restricted freedom and access to alternative sources of information, trust in government is (only) negatively correlated with the degree to which the country is a liberal democracy. Similarly, in societies where education is more limited, trust in government is less correlated with expert-based measures of performance.

Following the model proposed by Van Ryzin (2007[35]), satisfaction with services should exhibit a high correlation with objective or experience-based indicators despite being affected by external factors. Taking the results reported by Norris (2022[36]), in the context of OECD countries - with literate citizens and users who have access to a variety of sources of information about the performance of their government (from newspapers, social media, performance reports by their governments, etc.) -, even citizens who do not have recent experience with public services, should be able to provide an accurate account of the performance of services in their country.

## Methodology

This section seeks to test the relationship between satisfaction and more objective measures of service performance, drawing from the selection of indicators used in a chapter of the OECD *Government at a Glance* entitled 'Serving Citizens' to measure access, responsiveness and quality of health care, education and justice since 2015.

The first step was compiling a dataset with 67 indicators (shown in Table 3) for 37 OECD countries covering the period 2008-2019. The second step was to reduce the number of indicators, so they could be included in a regression. Several indicators were discarded due to time constraints, geographic coverage and/or small number of observations, or due to their high correlation with other indicators. As a result, 22 indicators from *Government at a Glance* were retained for 13 countries (Australia, Austria, the Czech Republic, Denmark, Finland, Hungary, Italy, Japan, the Netherlands, Norway, Slovenia, Spain, and Sweden), with a total of 156 observations for the years 2008-2019. These indicators are highlighted in bold in Table 3. The next step, was to employ factor analysis[14] to reduce these 22 indicators to their minimum common variance (i.e., decreasing the amount of indicators without losing information). Factor analysis produced five factors, without a significant loss of information (the five factors retained 80% of the common variance of these indicators).

Finally, the average satisfaction with health care, education and justice from Gallup World Poll's database was regressed against the resulting factors (which summarise 22 indicators of performance). While the Gallup World Poll relies on a small sample of respondents in each country (around 1,000 from main cities), and the questions only ask about an overall satisfaction judgement without controlling for biases, this is the only internationally comparable source across all OECD countries with a long time series.

[14] Factor analysis is a data reduction technique that reveals the relationship between observed variables and unobserved constructs (O'Leary-Kelly and Vokurka, 1998[45]). Factor analysis assumes that the covariance of observable variables can be explained (at least partially) by latent unobservable factors that share a common variance. There are two types of factor analysis, confirmatory factor analysis (CFA) and exploratory factor analysis (EFA). CFA is used to test a predefined model, which links certain observable variables to a single underlying construct. Instead, EFA does not rely on a predefined structure for the data. Hence, any observed variable could be correlated to any latent factor/dimension. This paper relies on EFA even though GAAG classifies each indicator into a specific factor/dimension (i.e. access, responsiveness, quality) because it retains most of the variance of the variables included in the analyses while reducing them to a smaller number of factors that can be correlated with satisfaction with services. The results of this analysis do not provide an indication of the reallocation of indicators among dimensions, since there is an imbalance between the services in the resulting factors, and the indicators themselves are not directly comparable between sectors.

## Table 3. Indicators used in *Government at a Glance* to measure the factors of the Serving Citizens Framework

| | Health care | | | Education | | | Justice | | |
|---|---|---|---|---|---|---|---|---|---|
| | Indicator | Expected Sign | Source | Indicator | Expected Sign | Source | Indicator | Expected Sign | Source |
| Access | Percentage of the population experiencing unmet care needs due to cost, distance or waiting times | - | EU-SILC | Share of private expenditures on education (primary to tertiary) | - | OECD Education at a Glance | **People can access and afford civil justice** | + | World Justice Project |
| | **Out-of-pocket medical expenditure as a share of final household expenditures** | - | OECD Health Statistics | **Enrolment rate at age 3** | + | OECD Education at a Glance | **Alternative dispute resolution mechanisms are accessible, impartial and effective** | + | World Justice Project |
| | Out-of-pocket payments as a share of total health spending | - | OECD Health Statistics | **Enrolment rate at age 4** | + | OECD Education at a Glance | Percentage of individuals who received legal advice | + | World Justice Project General Population Poll |
| | **Practising physicians per 1k population** | + | OECD Health Statistics | First-time tertiary entry rate for students under 25 years old | + | OECD Education at a Glance | Percentage of individuals who took action to solve their disputes | + | World Justice Project General Population Poll |
| | **Health care coverage** | + | OECD Health Statistics | **First time tertiary entry rate** | + | OECD Education at a Glance | Cost of enforcing contracts | - | World Bank's Ease of doing business |
| | Doctor visits | + | OECD Health Statistics | Percentage of students who have a computer to do homework at home | + | PISA | | | |
| Responsiveness | Median waiting time for hip replacement (from specialist assessment to treatment) | - | OECD Health Statistics | Percentage of early leavers from education and training aged 18-24 years who are not currently working | - | EU-LFS | Disposition time civil and commercial non-litigious cases, 1st instance | - | CEPEJ |
| | Median waiting time for cataract surgery (from specialist assessment to treatment) | - | OECD Health Statistics | **Percentage of young people (aged 15-29) not in education, employment or training** | - | OECD Education at a Glance | Disposition time civil and commercial litigious cases, 1st instance | - | CEPEJ |
| | Percentage of patients who got same or next-day appointment with doctor or nurse the last time they needed care | + | Commonwealth Fund International Policy Survey | Index of shortage of educational material | - | PISA | Disposition time administrative cases, 1st instance | - | CEPEJ |
| | Percentage of patients who waited two months or more for a specialist appointment | - | Commonwealth Fund International Policy Survey | Index of shortage of education staff | - | PISA | **Time for enforcing contracts** | - | World Bank's Ease of doing business |
| | Percentage of patients whose doctor often | + | Commonwealth Fund | % of students attending schools with | + | PISA | Number of procedures for | - | World Bank's |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | or always explains things in a way that is easy to understand | | International Policy Survey | rooms where homework can be done | | | enforcing contracts | | Ease of doing business |
| | Percentage of patients whose regular doctor involves them as much as they want in decisions about treatment or care | + | Commonwealth Fund International Policy Survey | % of students attending schools where staff provides help with homework | + | PISA | | | |
| | | | | **% of time lower secondary school principals spend talking to parents** | + | TALIS | | | |
| | | | | % of teachers reporting need for training in teaching students with special needs | - | TALIS | | | |
| Quality | **Mortality rate of acute myocardial infarction (heart attack)** | - | OECD Health Statistics | Mean score in reading | + | PISA | Effective enforcement of civil justice | + | World Justice Project |
| | **Mortality rate cerebrovascular disease (stroke)** | - | OECD Health Statistics | Mean score in mathematics | + | PISA | **Civil justice is free from improper government influence** | + | World Justice Project |
| | Breast cancer mortality in women | - | OECD Health Statistics | **Mean score in science** | + | PISA | **People do not use violence in response to personal grievances** | + | World Justice Project |
| | **30 day mortality rate following AMI hospitalisation** | - | OECD Health Statistics | Share of students below level 2 proficiency in reading | - | PISA | Crime is effectively controlled | + | World Justice Project |
| | **30 day mortality rate following ischemic stroke hospitalisation** | - | OECD Health Statistics | Share of students below level 2 proficiency in mathematics | - | PISA | Criminal justice system is timely and effective | + | World Justice Project |
| | **Asthma hospital admissions per 100k population** | - | OECD Health Statistics | **Share of students below level 2 proficiency in science** | - | PISA | | | |
| | **CHF hospital admissions per 100k population** | - | OECD Health Statistics | **% variation in reading performance explained by socioeconomic background** | - | PISA | | | |
| | COPD hospital admissions per 100k population | - | OECD Health Statistics | % variation in mathematics performance explained by socioeconomic background | - | PISA | | | |
| | **Diabetes hospital admissions per 100k population** | - | OECD Health Statistics | % variation in science performance explained by socioeconomic background | - | PISA | | | |
| | 5-year breast cancer survival rate (all stages) | + | OECD Health Statistics | Mean score in reading | + | PIRLS | | | |
| | | | | Index of cognitive adaptability | + | PISA | | | |
| | | | | Index of self-efficacy regarding global | + | PISA | | | |

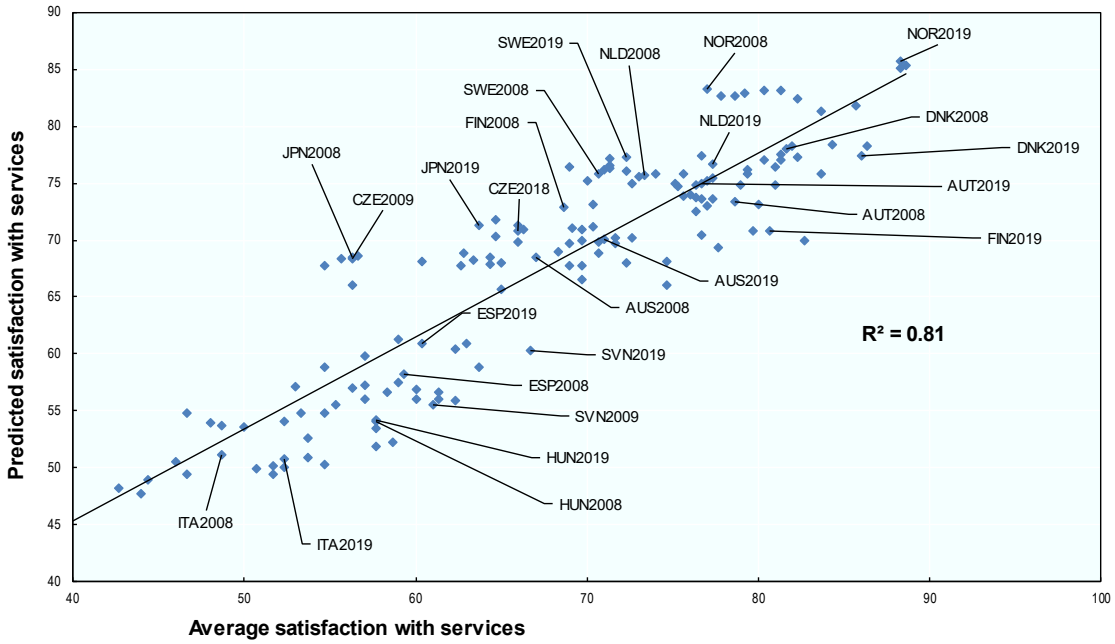
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | issues | | | | | |
| | | | | Mean score in mathematics (4th grade) | + | TIMSS | | | |
| | | | | Mean score in mathematics (8th grade) | + | TIMSS | | | |
| | | | | Mean score in science (4th grade) | + | TIMSS | | | |
| | | | | Mean score in science (8th grade) | + | TIMSS | | | |

Note: The table details all the indicators used in *Government at a Glance* since 2011, when a first chapter on measuring the performance of services was introduced. The classification of indicators into services and factors, together with their expected loadings come from the subsequent editions of *Government at a Glance* where they were used. A positive sign (+) indicates that an increase in the indicator (in its current form) is expected to contribute to an increase in the dimension and thus greater satisfaction with services. A negative sign (-) indicates that an increase in the indicator is expected to contribute to a decrease in the factor (e.g. access) and to lower satisfaction with services. The 22 indicators retained are in bold. Enrolment rates at ages 3 and 4 were averaged.

## Results

A regression between the five factors and average satisfaction with health care, education and justice, has an $R^2$ of 0.81. In other words, **81% of the change in satisfaction across countries and years can be associated with a change in the objective performance of services as measured by the 22 indicators from *Government at a Glance***. Figure 2 shows a good fit between predicted satisfaction (using the results of the regression) and observed satisfaction. Annex A provides more details about the results of the regression.

### Figure 2. Observed and predicted satisfaction with services




Note: The predicted satisfaction with services is computed using the coefficients of the regression between the rotated factor scores (which come from aggregating the indicators in bold in Table 3 into 5 factors via factor analysis) and the average satisfaction with education, health care and justice using Gallup World Poll.
Source: Average satisfaction with services comes from Gallup World Poll.

For some countries, such as Austria, the predicted satisfaction is lower than the observed one (on average, 3 percentage points less), and the opposite happens for others, such as the Czech Republic (predicted satisfaction is on average 8 percentage points higher). The change in satisfaction that is not explained by a change in the performance of services using these 22 indicators can be attributable to external factors, such as biases in the design of the questionnaire, cognitive biases from respondents (e.g. negativity bias, "anti-public sector bias") and/or differing expectations across countries (which can affect satisfaction independently of citizen's actual experience with services). There are also other underlying, country-specific differences that are unrelated to service delivery or measurement, such as culture, socioeconomic differences, etc.

The frequency of data collection is another factor explaining the differences between the change in objective performance and the change in satisfaction. Satisfaction data are produced annually, while the other indicators are produced at less frequent intervals (which required imputing data to have complete observations for each country/year); hence satisfaction is more likely than the model to reflect "jumps" in performance. For example, if doctors and nurses were to go on strike it would certainly affect access and

responsiveness of health care, yet, the drop in performance would not be captured by the model due to the unfrequent and lengthy data collection process of the indicators measuring these dimensions (e.g. compiling waiting times for elective surgery from administrative sources requires standardising and validating records across all hospitals in a given country).

These results show that, generally speaking, satisfaction positively correlates with the performance of services, as measured by objective and experience-based indicators selected by sector experts. In other words, satisfaction, when measured well, can help decision-makers in identifying trends in aggregate service performance.

Nevertheless, at least in cross-national evaluations such as the one presented in this paper, overall satisfaction measures remain an aggregate indicator, summarising a wide range of aspects of performance, which need to be unpacked in order to identify areas that need improvement, and to take concrete actions to address performance gaps. The next section discusses how governments use service satisfaction data and objective measures to inform service design and delivery.

# Utilising users' experience and service performance data to improve public services

As shown in the first section, an increasing number of OECD national/federal governments consider satisfaction with public services as a goal of high-level policy agendas (e.g. US President's "Management Agenda" or Australia's "Delivering for Australians" reform agenda). They then appoint a central agency to collect data and report on key indicators of performance, in addition to the performance systems set up by service delivery agencies and line ministries. However, as seen in the previous section, using satisfaction data in decision-making is not straightforward and needs to be considered alongside other indicators.

This section builds on the results of the workshop "Serving citizens: measuring the performance of services for better delivery" and experiences presented by country experts during the workshop to start identifying a number of key issues in measuring performance of public services and making effective use of this evidence. The aim is to build on past experience and good practices to guide governments in identifying key building blocks to enhance citizens' and users' experience and satisfaction with public services. As a summary for further analysis, this paper proposes to consider the following aspects as preliminary guidance on improving the use of performance data in decision making:

- Seeking support from the political leadership to the improvement of services from a citizens' and users' perspective,
- Identifying one or several co-ordinated agencies who should be responsible for monitoring the performance of services against citizens' and users' needs and expectations and who they are accountable to,
- Designing a monitoring system including satisfaction indicators alongside other more objective measures of the performance of services,
- Ensuring cooperation between agencies to improve services for citizens and users.

## Support from the political leadership

Improving public services from a citizens' and users' perspective (focusing on their needs and the aspects they value with regard to public services) is impossible without support from the political leadership.

This support can translate into a national policy on the improvement of public services that includes a commitment to focus on users' needs, and clear plans on how the performance of public agencies will be monitored (e.g. what indicators will be used, what will be the frequency of monitoring, etc.). This includes

explaining how the monitoring of results will be disseminated to citizens and users, and used in the public debate, in decision-making and in auditing processes (transparency and accountability).

The cases presented in this paper, from Australia, France, Ireland and the United States are examples of how government support of service improvement has translated into medium-term policies that are closely linked to government priorities. Support from the political leadership is the basis for the co-operation between agencies, which is required to move past the silos and legacy processes/regulations of the public administration to focus on citizens and users and their needs. This usually involves overhauling services from the user perspective (e.g. the United States' Executive Order on Transforming Federal Customer Experience and Service Delivery to Rebuild Trust in Government from 13 December 2021 or France's Discours du Premier ministre au Comité interministériel de la transformation publique from 5 February 2021).

Political support is also needed to ensure that the objective of improving services is pursued beyond the achievement of measurable indicators and targets. This includes incentivising government agencies to ensure that citizens and users have a seamless experience with all government services. Adopting a user journey or life-events approach is one such example. Ideally governments would adopt a "government as a platform" approach to the design and delivery of services. This entails sharing best practices and guidelines (e.g. how to write in plain language), as well as common reusable components (e.g. digital identity, common hosting and infrastructure, etc.) – for which there would be central investment to reduce the burden on service delivery agencies that have small budgets –, human resources (e.g. training, recruitment, transfer of staff, internal consulting teams, etc.), and governance arrangements (e.g. procurement systems and quality assurance mechanisms), among other aspects (OECD, 2020[37]). This would also help in providing an integrated experience to citizens and users across levels of government.

## Accountability for monitoring service performance and results

The responsibility of monitoring the performance of user journeys and life events and for pursuing the objective of improving public services from a citizens' and users' perspective in general should fall on one or more co-ordinated agencies. The agency (or agencies) responsible for embedding a citizen and user perspective in the public administration would typically have tasks such as ensuring support from the political leadership and from the agencies responsible for the implementation of the policy, co-ordinating among these institutions. This is especially true when they are co-responsible for a life event or user journey or contribute to achieving each other's objectives. This is achieved by simplifying interaction with users' and citizens' (e.g. school vaccination clinics to increase child immunization by leveraging on the opportunity of having them at school every day) and reporting on results. In particular, there should be an explicit mention of who this/these agency(ies) are accountable to (e.g. to citizens and users, to the parliament, to the head of government), and how often they are expected to report on results. This adds to the existing sectorial policies and objectives, which require specialised expert knowledge to be achieved (e.g. the ministry of education is obviously in an excellent position to monitor student learning).

A wide range of mechanisms are available to support the goals of improving public services and ensure follow up. These include having regular formal and/or informal discussions about performance results (e.g. survey results) with high-level civil servants from the agencies responsible for the delivery of a service or journey or with external stakeholders (e.g. academia or the private sector). Mechanisms could also include high level discussions with the heads of the agencies that are responsible for a service or a user journey. In Ireland, the Quality Customer Service Network, in which Quality Customer Service Officers (responsible for overseeing customer service standards in their organisations) participate, facilitates the sharing of best practices in service delivery across the government in a more informal setting.

As part of the co-ordination between agencies, many countries provide guidance to service providers. This guidance can include how to identify capacity gaps, or how to develop action plans to translate the results

of performance monitoring into improvement of public services. For example, the OMB in the United States provides guidance to High Impact Service Providers to analyse their capacity to deliver services (e.g. capacity to measure outcomes, governance and strategy, customer understanding) and to develop annual action plans to improve such capacity. Ireland's Public Sector Reform office provides guidance to organisations on developing customer charters and designing action plans to put them into practice, as well as on implementing the innovation strategy, which aims at "delivering world-class public services in Ireland". Similarly, France's DITP advises service providers to implement a continuous improvement approach.

Matching performance data and analysis with the needs and timing of policy-making processes is a challenge encountered in many countries, which can undermine the accountability of the delivery agencies. To remedy this issue, some countries (e.g. United States) have established a series of processes to ensure that performance management (including capacity assessments) is integrated in the annual planning of service providers (see OMB Circular A-11 Section 280).

Reporting results to citizens and users is a key feature of the policies to improve public services for two main reasons. First, this strengthens the accountability of public agencies for delivering results, and enhancing public administration transparency. In addition, some studies have found that citizens are more satisfied if services meet their positive expectations. Positive expectations can be shaped by accessing information about the good performance of services (e.g. (Grimmelikhuijsen and Porumbescu, 2017[26]), (Van Ryzin, 2013[38])). France's DITP[15], the United States' OMB[16] and UK's Government Digital Service [17] have developed portals where citizens can consult key metrics of public service performance (e.g. satisfaction with services, wait times to receive answers, costs per transaction, etc.).

## Choice of indicators and monitoring systems

A performance monitoring system should have a distinct user focus, which entails being clear about what characteristics users (of the service and of performance information) value and the standards they expect (Pidd, 2012[39]). 'Value' involves what matters to individuals as well as the wider society or institution. Most public agencies at the central/federal level have a dual monitoring system. One system is specific to them, to track the achievement of their own key objectives for which they have expert knowledge (e.g. increasing tax compliance and enhancing tax collection for the tax authorities). The other is related to the experience they deliver to citizens and users, for which they need to cooperate with other agencies (e.g. the tax administration's would need to cooperate with the motor-vehicle registration in the journey of selling a used car to ensure that citizens and companies do not need to provide the same information several times). These two monitoring systems are not mutually exclusive, as long as their contribution to the work of each agency responsible for delivering services is clear.

In addition, decision-makers need timely data that is simple to interpret and to draw insights from to make decisions, given that they are responsible for a wide range of topics that compete for their attention.

Satisfaction is one of the most common indicators of service performance because it reflects citizens' and users' experience with services, and it is faster (and cheaper) to collect and disseminate than objective measures of performance. For this reason, satisfaction surveys have proliferated in the public administration. Most countries now have two types of surveys. One type is conducted by individual service

[15] https://www.plus.transformation.gouv.fr/recherche-de-resultats

[16] https://www.performance.gov/cx/

[17] https://data.gov.uk/dataset/731b25a8-0462-4a7d-aa3f-5a5d44ae26d2/historical-performance-platform

providers (e.g. the health insurance administration) and can focus on each individual service they provide. The second type is conducted by the central government (as shown in the example of Norway).

Surveys by individual service providers can be useful to describe the profile of users (e.g. socio-economic background) and their recent experience with the service. Typically, these surveys ask citizens and users about the accessibility (e.g. affordability, availability of channels, access to information and clarity of procedures), responsiveness (e.g. courtesy of staff and timeliness) and quality (e.g. quality of infrastructure, competence of staff and outcome) of services that they have used. These surveys target service users: thus, only individuals who have successfully accessed (and likely, used) the service are included. These surveys monitor issues that contribute to achieving the objectives of an individual agency (e.g. a health ministry that measures whether doctors provide sufficient explanations when patients have questions or doubts).

Central surveys can address the population in general and explore the barriers of access to services, identifying groups that struggle to have their needs met by the public administration. Central surveys can focus on user journeys or life events (e.g. losing the ability to work, getting married), asking citizens and users how they solved their needs (e.g. how many agencies they interacted with, what for, etc.). For example, France's "Baromètre de la Complexité" focuses on understanding how complex it was for users to resolve a problem (e.g. dealing with the death of a loved one) while interacting with public services, and identifying the most cumbersome aspects (e.g. the number of documents that are required to request a service, the type of language used in forms). These surveys help identify common barriers (e.g. use of jargon, complex processes) that could benefit from common solutions.

Yet, while these surveys can provide an accurate account of the trends in the performance of public services, overall satisfaction ratings summarise a wide range of aspects of service performance (e.g. accessibility, responsiveness and service quality) that need to be unpacked to inform decision-making. Satisfaction data can contribute to stimulating discussions about service performance trends and alert decision-makers to take action. However, without other indicators or sources of data, such data may not provide insight into what specific aspects are underperforming or what improvements are needed (e.g. if certain groups of the population are not able to access a service because it is not affordable for them, should they receive a subsidy? should fees be reduced? are there any other reasons why the service is so costly?).

A complete performance measurement system for public services should include process (e.g. waiting times in hospitals), output (e.g. survival rates), service quality (e.g. satisfaction with the service), and outcome (e.g. are people better off as a result of the service?) measures to give a balanced view of the performance of the institution, reflecting the most important aspects of the mission of an organisation (Pidd, 2012[39]). However, different sources of data (and indicators) can provide diverging pictures of the actual performance of the agency. Therefore, more time is needed to analyse and reconcile these sources of information and to draw insights that can help in improving services.

Usually, quantitative data allow governments to identify the areas where there is a problem that needs attention but are not the best when it comes to providing solutions, especially as they pertain to citizens' and users' experience with a service. Indeed, qualitative research provides a better understanding of the feelings, motivations and experiences of users, which can more easily provide insights into user pain points. It is important to bear in mind that while qualitative research methods and direct user feedback usually provide more actionable insights, quantitative data are needed to understand the extent to which the data coming from non-representative samples can be generalised to the whole population[18].

[18] Note that 'population' refers to a group under study, and this can be a sub-group (e.g. youth) of a country's population.

Transforming data into insight requires using service performance data from various sources (e.g. satisfaction surveys, user research, citizens' and users' feedback, objective performance indicators, etc.) to model the experience of a user with a service. For example, in France, the DITP has developed a "fiche" for each life event, detailing the user journey – using the results of the "Baromètre de la Complexité" in which users indicate which agencies they interacted with to solve a concern, such as claiming unemployment benefits, and other sources – and identifying pain points, or bottlenecks, to then resolve them. Some pain points may stem from the language used by the public administration in communicating with citizens and users (e.g. legal jargon), which make accessing information about the service difficult. Other solutions to analyse data can entail creating maps, dashboards or infographics.

## Co-operation between agencies to enhance the performance of services from a citizen's perspective.

In the countries with a central policy to improve services from a citizen point of view (following user design principles), there should be co-operation between the agency responsible for monitoring the performance of key life events and user journeys and the ministries or agencies directly responsible for their delivery. This co-operation is sensitive because it requires openness from the agencies directly responsible for the provision of public services to having conversations about potential changes in their area of competence with external actors. Sometimes these changes entail dismantling existing processes and changing regulations (and un-learning them), which may be difficult, especially in the short run.

Paying narrow attention to performance data can result in "tunnel vision" (i.e. focusing on quantifiable indicators at the expense of unquantifiable objectives). Most agencies have their own monitoring systems, which may provide results that differ from those computed by central agencies. For example, a survey sent to users of a service by its provider (e.g. after completing a transaction) may indicate that most users are satisfied, while a survey sent to citizens and users in general may indicate otherwise. As a result, the discussion between the central agency and the service provider may end up focusing only on which of the two results is correct. This tunnel vision does not contribute to the aim of improving services. In fact, confirmation bias (the tendency to favour information that supports one's previous beliefs) may prevent relevant stakeholders from reaching an understanding that contributes to improving the performance of services.

Instead of identifying good or bad performers and singling them out, emphasising performance aspects that are common to all service providers (or at least some of them) contributes to enhancing co-operation. For example, performance gaps may be identified among population segments or for a specific life event that requires the intervention of multiple service providers. When this occurs, improving the performance of services requires strengthening the co-operation between service providers to break their silos and act in a co-ordinated manner. In Australia, the Citizen Engagement Unit from the Department of the Prime Minister and Cabinet (PM&C) identifies key "hotspots" that require further research and improvement. Box 4 describes how the results of the Citizen Engagement Survey allowed the country to improve services for young people.

### Box 5. From insight to impact: improving young people's satisfaction with public services in Australia

In 2020, Australia's Citizen Experience Survey indicated young people (18-34) were having difficulty accessing public services and had low levels of trust and satisfaction in public service. The Department commissioned Monash University to conduct an in-depth analysis of young people's experiences accessing public services. Citizen Experience Survey results were supplemented with qualitative and quantitative data as well as research that included focus groups with young people across Australia.

The results from their research, *The Young Adults' Experiences with Australian Public Services Report,* described young people's pathways into adulthood and how they engage with public services along the way. Nine specific areas of focus were identified as opportunities for further action to improve the service delivery experiences of young people, including improving access to information about public services.

The results informed Australia's Youth Policy Framework, which was released in 2021. The Framework sets out a blueprint for how the Government is supporting young people to navigate life's challenges and remain engaged active citizens who can contribute productively to society.

Note: Australian public services refer to federal, national or commonwealth public service
Source: (Monash University, 2020[40])

Focusing – at all times – on citizens and users can help in identifying solutions and mobilising stakeholders within the public administration to implement them. For example, in 2010, the UK put in place the "tell us once" service, by which at the time of registering the death of an individual, all other relevant government organisations are informed at once. This initiative was launched by the Department for Work and Pensions in co-operation with central and local government, and was motivated by the feedback of a user who struggled to get through all the paperwork required to notify authorities of the passing of a loved one (Briggs, 2011[41]). This example illustrates the concept of citizen-centricity in practice: gathering feedback from users – whose lives are impacted by public services – through qualitative and quantitative research. Another source of feedback for improving public services includes suggestions from people's own experience, especially people who struggle to have their needs met.

Initiatives that break silos in the public administration to focus on users' needs usually face numerous roadblocks (e.g. data protection regulations, lack of funding) that can only be surmounted with strong commitment from those involved.

# Conclusion and next steps

Our societies are changing fast, driven by population ageing, climate change adaptation, digitalisation, and the changing world of work, among other trends. Governments face the challenge of investing to adapt to these trends, while continuing to ensure that public services remain accessible, responsive and of good quality. This is crucial, not only because services address a wide variety of individual and social needs (from health care and education to safety), but also because the performance of public services affects citizens' assessment of their government and their attitudes towards public institutions.

Satisfaction with services has become a widely used indicator to measure the performance of public services from a citizens' and users' perspective. Satisfaction is included as a key measure in monitoring the achievement of high-level strategies at the national and international level, such as the Sustainable Development Goals. However, satisfaction is a subjective measure, and it has several limitations.

This paper discusses how to enhance the validity of satisfaction measures in monitoring the performance of public services, and shows that - for OECD countries - satisfaction is positively correlated with the performance of services across countries and years at the macro level. As such, it complements performance information from administrative sources and objective indicators.

Finally, this paper discusses countries' practices in designing and using public service performance evidence to improve users' experience. It presents several challenges found throughout OECD countries, such as the role of high-level political support for performance measurement systems, co-operation across public administrations, choice of metrics in the monitoring systems, accountability, etc.

Making use of the good practices shared by a number of OECD countries, the paper provides some preliminary guidance on improving the use of performance data in decision making. This guidance is not complete and will benefit from further exchanges of practices and standards from public administration practitioners across OECD countries.

# References


Andersen, S. and M. Hjortskov (2015), "Cognitive Biases in Performance Evaluations", *Journal of Public Administration Research and Theory*, Vol. 26/4, pp. 647–662, https://doi.org/10.1093/jopart/muv036. [25]

Andrews, R., G. Boyne and R. Walker (2011), "The Impact of Management on Administrative and Survey Measures of Organizational Performance", *Public Management Review*, pp. 227-255, https://doi.org/10.1080/14719037.2010.532968. [28]

Andrews, R., G. Boyne and R. Walker (eds.) (2006), *Subjective and objective measures of organizational performance : An empirical exploration*, Cambridge University Press, https://doi.org/10.1017/CBO9780511488511.002. [16]

Australian Government (2019), *Delivering for Australians: A world-class Australian Public Service: The Government's APS reform agenda*. [4]

Brezzi, M. et al. (2021), "An updated OECD framework on drivers of trust in public institutions to meet current and future challenges", *OECD Working Papers on Public Governance*, No. 48, OECD Publishing, Paris, https://doi.org/10.1787/b6c5478c-en. [50]

Briggs, M. (2011), *The benefits of Tell Us Once*, https://www.theguardian.com/government-computing-network/2011/dec/16/tell-us-once-matt-briggs. [41]

Dinno, A. (2001), "PARAN: Stata module to compute Horn's test of principal components/factors", *Statistical Software Components S420702, Boston College Department of Economics,*, p. S420702. [47]

Direction interministérielle de la Transformation publique (2021), *Services Publics+*. [8]

Executive Office of Management and Budget (2021), *Preparation, Submission, and Execution of the Budget,Circular no. A–11 § 280*. [6]

Frey, V. (2021), *Did bigger and better benefits during the pandemic ease people's money worries?*, https://www.washingtonpost.com/politics/2021/08/25/did-bigger-better-benefits-during-pandemic-ease-peoples-money-worries/. [32]

González, S. (2020), "Testing the evidence, how good are public sector responsiveness measures and how to improve them?", *OECD Working Papers on Public Governance*, https://doi.org/doi.org/10.1787/c1b10334-en. [46]

Goodman, E., M. Kuniavsky and A. Moed (2012), *Observing the User Experience: A Practitioner's Guide to User Research*, Morgan Kaufmann, https://doi.org/10.1016/C2010-0-64844-9. [14]

Grimmelikhuijsen, S. and G. Porumbescu (2017), "Reconsidering the expectancy disconfirmation model. Three experimental replications", *Public Management Review*, Vol. 19/9, pp. 1272-1292, https://doi.org/10.1080/14719037.2017.1282000. [26]

Howard, C. (2010), "Are we being served? A critical perspective on Canada's Citizens First satisfaction surveys", *International Review of Administrative Sciences*, Vol. 76/1, pp. 65–83, https://doi.org/10.1177/0020852309359045. [27]

Ipsos-MORI (2006), *Understanding Public and Patient Attitudes to the NHS*, https://www.ipsos.com/sites/default/files/migrations/en-uk/files/Assets/Docs/Archive/Polls/healthcarecommission.pdf. [21]

Ireland's Department of Public Expenditure and Reform (2017), *Our Public Service 2020*. [7]

James, O. (2011), "1. Managing citizens' expectations of public service performance: evidence from observation and experimentation in localgovernment", *Public Administration*, Vol. 89/4, pp. 1419–1435, https://doi.org/10.1111/j.1467-9299.2011.01962.x. [24]

Kelley, E. and J. Hurst (2006), "Health Care Quality Indicators Project: Conceptual Framework Paper", *OECD Health Working Papers*, No. 23, OECD Publishing, Paris, https://doi.org/10.1787/440134737301. [49]

Ladhari, R. (2009), "Service quality, emotional satisfaction, and behavioural intentions: A study in the hotel industry", *Managing Service Quality: An International Journal*, Vol. 19/3, pp. 308-331, https://doi.org/10.1108/09604520910955320. [44]

Levesque, J., M. Harris and G. Russell (2013), "Patient-centred access to health care: conceptualising access at the interface of health systems and populations", *International Journal for Equity in Health*, Vol. 12/18, https://doi.org/10.1186/1475-9276-12-18. [43]

Li, Y., L. Ren and F. Luo (2016), "Is bad stronger than good? The impact of police-citizen encounters on public satisfaction with police", *Policing: An International Journal*, Vol. 39/1, pp. 109-126, https://doi.org/10.1108/PIJPSM-05-2015-0058. [22]

Marvel, J. (2015), "Public Opinion and Public Sector Performance: Are Individuals' Beliefs About Performance Evidence-Based or the Product of Anti–Public Sector Bias?", *International Public Management Journal*, Vol. 18/2, pp. 209-227, https://doi.org/10.1080/10967494.2014.996627. [20]

Monash University (2020), *Young adults' experiences with Australian public services*, https://www.monash.edu/arts/social-sciences/australians-experiences-with-public-services/research-projects/young-adults-experience-with-australian-public-services2. [40]

Norris, P. (2022), "Trust in Government Redux: The Role of Information Environments and Cognitive Skills", *HKS Faculty Research Working Paper Series RWP22-001*, https://www.hks.harvard.edu/publications/trust-government-redux-role-information-environments-and-cognitive-skills. [36]

OECD (2021), *Drivers of Trust in Public Institutions in Finland*, OECD Publishing, Paris, https://doi.org/10.1787/52600c9e-en. [1]

OECD (2020), *Digital Government in Chile – Improving Public Service Design and Delivery, OECD Digital Government Studies*, OECD Publishing, https://doi.org/10.1787/b94582e8-en. [37]

OECD (2020), "Digital Government Index: 2019 results", *OECD Public Governance Policy Papers*, No. 03, OECD Publishing, Paris, https://doi.org/10.1787/4de9f5bb-en. [10]

OECD (2020), "The OECD Digital Government Policy Framework: Six Dimensions of a Digital Government", *OECD Public Governance Policy Papers*, Vol. 2, https://doi.org/10.1787/f64fed2a-en. [9]

OECD (2019), *Government at a Glance 2019*, OECD Publishing, Paris, https://doi.org/10.1787/8ccf5c38-en. [48]

OECD (2019), *Main Findings from the 2018 Risks that Matter Survey*, OECD Publishing, Paris, https://doi.org/10.1787/9266e48a-en. [31]

OECD (2015), *Recommendation of the OECD Council on good statistical practice*, https://www.oecd.org/statistics/good-practice-toolkit/Brochure-Good-Stat-Practices.pdf. [15]

O'Leary-Kelly, S. and R. Vokurka (1998), "The Empirical Assessment of Construct Validity", *Journal of Operations Management*, Vol. 16/4, pp. 387-405, https://doi.org/10.1016/S0272-6963(98)00020-5. [45]

Olsen, A. (2015), "Citizen (Dis)satisfaction: An Experimental Equivalence Framing Study", *Public Administration Review*, Vol. 75/3, pp. 469-478, https://doi.org/10.1111/puar.12337. [23]

Ongaro, E. and S. Van Thiel (eds.) (2018), *Explaining Citizen Satisfaction and Dissatisfaction with Public Services*, Palgrave Macmillan, https://doi.org/10.1057/978-1-137-55269-3_11. [11]

OPSI and MBRCGI (2020), *Embracing Innovation in Government: Innovative responses to the COVID-19 crisis*, https://trends.oecd-opsi.org/wp-content/uploads/2020/11/OECD-Innovative-Responses-to-Covid-19.pdf. [51]

Parks, R. (1984), "Linking Objective and Subjective Measures of Performance", *Public Administration Review*, Vol. 44/2, pp. 118-127, https://www.jstor.org/stable/975862. [33]

Peeters, R., A. Gofen and O. Meza (2020), "Gaming the system: Responses to dissatisfaction with public services beyond exit and voice", *Public administration*, Vol. 98/4, pp. 824-839, https://doi.org/10.1111/padm.12680. [3]

Pidd, M. (2012), *Measuring the Performance of Public Services: Principles and Practice*, Cambridge University Press, https://doi.org/10.1017/CBO9780511791550. [39]

Poister, T. and J. Thomas (2011), "The Effect of Expectations and Expectancy Confirmation/Disconfirmation on Motorists' Satisfaction with State Highways", *Journal of Public Administration Research and Theory*, Vol. 21/4, pp. 601-617, https://doi.org/10.2307/41342597. [18]

Praia City Group (2020), *Handbook on Governance Statistics*, Instituto Nacional de Estatística, https://ine.cv/praiagroup/handbook-on-governance-statistics/. [13]

Republic of Slovenia (2016), *Digital Slovenia 2020 - Development strategy for the information society until 2020*. [12]

Stipak, B. (1979), "Satisfaction with Urban Services: Potential Misuse as a Performance Indicator", *Public Administration Review*, Vol. 39/1, pp. 46-52, http://www.jstor.com/stable/3110378. [29]

Swindell, D. and J. Kelly (2002), "A Multiple-Indicator Approach to Municipal Service Evaluation: Correlating Performance Measurement and Citizen Satisfaction across Jurisdictions", *Public Administration Review*, pp. 610 - 621. [30]

UNDP Oslo Governance Centre (2020), *Indicator 16.6.2: Proportion of population satisfied with their last experience of public services*, https://unstats.un.org/sdgs/metadata/files/Metadata-16-06-02.pdf. [19]

United States' President's Management Council (2021), *The Biden-Harris Management Agenda Vision: a roadmap for our Government to deliver results for all Americans*. [5]

Van de Walle, S. and G. Bouckaert (2003), "Public Service Performance and Trust in Government: The Problem of Causality", *International Journal of Public Administration*, Vol. 26/8-9, pp. 891-913, https://doi.org/10.1081/PAD-120019352. [2]

Van Ryzin, G. (2013), "An Experimental Test of the Expectancy-Disconfirmation Theory of Citizen Satisfaction", *Journal of Policy Analysis and Management*, Vol. 32/3, pp. 597-614, https://doi.org/10.1002/pam.21702. [38]

Van Ryzin, G. (2007), "Pieces of a Puzzle: Linking Government", *Public Performance & Management Review*, Vol. 30/4, pp. 521-535, https://doi.org/10.2753/PMR1530-9576300403. [35]

Van Ryzin, G. (2004), "Expectations, performance, and citizen satisfaction with urban services", *Journal of Policy Analysis and Management*, Vol. 23/3, pp. 433-448, https://doi.org/10.1002/pam.20020. [17]

Van Ryzin, G., S. Immerwahr and S. Altman (2008), "Measuring Street Cleanliness: A Comparison of New York City's Scorecard and Results from a Citizen Survey", *Public Administration Review*, Vol. 68/2, pp. 295-303, https://doi.org/10.1111/j.1540-6210.2007.00863.x. [34]

Yang, K. and M. Holzer (2006), "The Performance–Trust Link: Implications for Performance Measurement", *Public Administration Review*, Vol. 66/1, pp. 114-126, https://doi.org/10.1111/j.1540-6210.2006.00560.x. [42]

# Annexe A. Methodology and results

## Dataset and data imputation

A dataset containing the 67 indicators from Table 3 for 37 OECD countries between 2008 and 2019 was compiled. In cases of missing data, two types of imputation were used. For missing values between two observations (e.g. an indicator was measured in 2015 and 2018, hence data for 2016 and 2017 are missing), a weighted mean was computed, assigning a larger weight to the closest observation (e.g. for 2016, the weights would be 2/3 for 2015 and 1/3 for 2018). For missing values before the first or after the last observation, the mean value for the country was imputed.

### *Methodology*

In order to determine the appropriate number of factors, the following steps were followed:

1. Determine if at least some indicators are correlated (Pearson's r > 0.5 or <-0.5), and discard at least one of those that are almost linearly correlated (r > 0.90)
2. Transform all indicators into z-scores (observed value minus mean for the indicator divided by its standard deviation)
3. Compute the Kaiser-Meyer-Olkin (KMO) test of sampling adequacy for all relevant indicators, Discard those with a KMO of less than 0.5.
4. Run factor analysis, and discard those indicators with a commonality of less than 0.5.
5. Discard indicators that have an item-rest correlation of less than 0.1. Item rest correlation is the Pearson correlation coefficient between an indicator and the index composed of the rest of the indicators.
6. Determine the number of factors to keep by: (a) analysing the Scree plot and keeping only the factors above the "elbow", (b) running parallel analysis, which creates a random dataset with the same number of variables and observations as the original data, and it computes the Eigenvalues derived from the random data. If the Eigenvalues from the original dataset are lower than the ones from the random one, then factor should be discarded, (c) applying the Kaiser criterion (only factors with an Eigenvalue of 1 or above).
7. Perform a Promax factor rotation (oblique), which allows factors to be correlated, since there are indicators that have large loadings for more than one factor.
8. Compute the Cronbach's alpha coefficient for retained indicators. Only solutions with a coefficient higher than 0.7 were kept for the analysis.
9. Analyse goodness of fit with the Akaike Information Criterion (AIC) and the Bayes Information Criterion (BIC), choosing the model with the smallest value in each of them.
10. Predict factor rotated factor scores for each country and year by means of regression.

### *Results*

A first step, before conducting factor analysis, was to explore the correlation between the indicators used to measure the performance of each service (education, health and justice)[19]. Analysing the correlation structure of the data is a necessary step to conduct factor analysis: if there is no correlation between indicators, no factors can emerge. Some indicators that are highly correlated with others were discarded from the analysis. Figures A.3, A.4 and A.5 show the correlation matrix by service.

## Figure A.1. Correlation matrix of health indicators

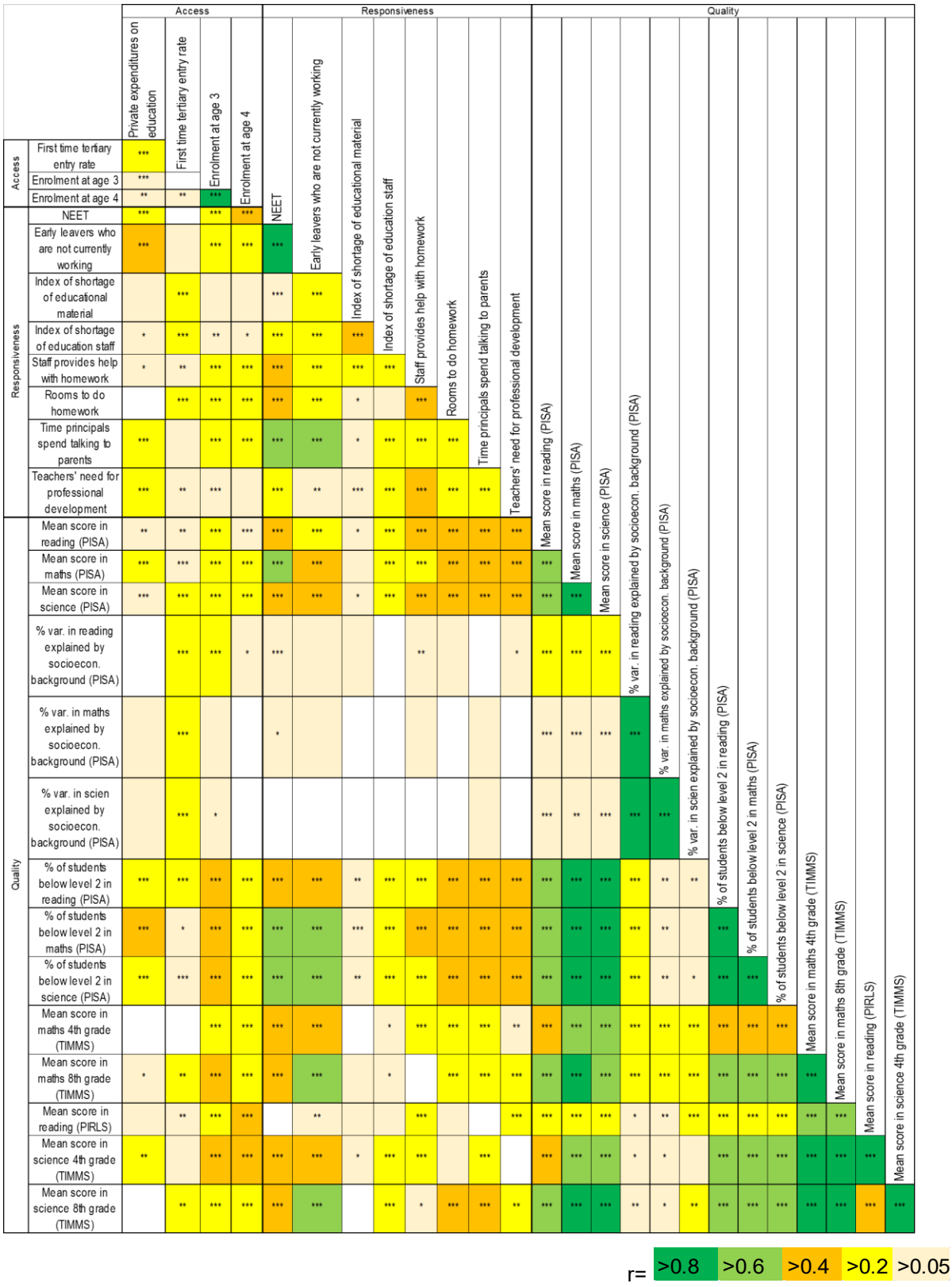
| | | Access | | | | | Responsiveness | | | Quality | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Health care coverage | Unmet needs for health care | Outofpocket spending | Physicians per 100k population | Average of doctor visits | Median waiting time for hip replacement | Median waiting time for knee replacement | Median waiting time for cataract surgery | CHF hospital admissions | Asthma hospital admissions | COPD hospital admissions | Diabetes hospital admissions | 30 day mortality following stroke hospitalisation | 30 day mortality following AMI hospitalisation | Mortality rate of AMI | Mortality rate stroke |
| Acces | Unmet needs for health care | *** | | | | | | | | | | | | | | | |
| | Out-of-pocket spending | | * | | | | | | | | | | | | | | |
| | Physicians per 100k population | *** | *** | *** | | | | | | | | | | | | | |
| | Average of doctor visits | ** | | | | | | | | | | | | | | | |
| Responsiveness | Median waiting time for hip replacement | *** | *** | * | *** | | | | | | | | | | | | |
| | Median waiting time for knee replacement | *** | *** | | ** | * | *** | | | | | | | | | | |
| | Median waiting time for cataract surgery | *** | *** | ** | *** | | *** | *** | | | | | | | | | |
| Quality | CHF hospital admissions | *** | | | ** | *** | | | *** | | | | | | | | |
| | Asthma hospital admissions | ** | *** | ** | *** | *** | | * | *** | *** | | | | | | | |
| | COPD hospital admissions | | | | ** | | *** | *** | * | * | *** | | | | | | |
| | Diabetes hospital admissions | *** | | | *** | *** | *** | *** | *** | *** | *** | * | | | | | |
| | 30 day mortality following stroke hospitalisation | *** | *** | | *** | *** | *** | *** | *** | *** | ** | | *** | | | | |
| | 30 day mortality following AMI hospitalisation | *** | *** | *** | *** | | *** | *** | | * | | *** | *** | *** | | | |
| | Mortality rate of AMI | *** | *** | | *** | *** | | | | | ** | *** | *** | *** | *** | | |
| | Mortality rate stroke | *** | *** | *** | | *** | * | *** | *** | *** | *** | | *** | *** | *** | *** | |
| | Breast cancer mortality in women | *** | *** | *** | *** | *** | *** | *** | * | *** | | *** | *** | ** | *** | *** | |

*r*= >0.8 >0.6 >0.4 >0.2 >0.05

Note: Significance levels: * p<0.05, ** p<0.01, *** p<0.001. The correlation coefficient (r) is indicated in absolute terms (modulus).

[19] The number of observations and of indicators included in the correlations was larger than the number of observations retained for factor analysis, ranging from 156 observations (for the correlation between TIMMS mathematics for 4th and World Justice Project indicators, e.g. people can access and afford civil justice) to 444 observations for indicators from PISA (e.g. index of shortage of educational material and rooms to do homework). These correlations included imputed data because of the spacing between observations for each indicator; nevertheless, the imputations do not necessarily affect the correlations between indicators because there is more variation across countries than across years for the observations.

## Figure A.2. Correlation matrix of education indicators




Note: Significance levels: * p<0.05, ** p<0.01, *** p<0.001. The correlation coefficient (r) is indicated in absolute terms (modulus).

## Figure A.3. Correlation matrix of justice indicators

| | | Access | | | | | Responsiveness | | | | | Quality | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | People can access and afford civil justice | ADRs are accessible, impartial, and effective | Sought legal advise | Took action to solve their dispute | Cost of enforcing contracts | Processes (number) for enforcing c | DT administrative cases | Disposition time civil and commercial litigio | Disposition time civil, commercial and administ | Time for enforcing contracts | Crime is effectively controlled | People do not resort to violence | Civil justice is free of improper government influence | Civil justice is effectively enforced |
| Access | ADRs are accessible, impartial, and effective | *** | | | | | | | | | | | | | |
| | Sought legal advise | *** | ** | | | | | | | | | | | | |
| | Took action to solve their dispute | | *** | *** | | | | | | | | | | | |
| | Cost of enforcing contracts | *** | *** | ** | | | | | | | | | | | |
| Responsiveness | Processes (number) for enforcing contracts | * | *** | | *** | | | | | | | | | | |
| | DT administrative cases | | | *** | | | *** | | | | | | | | |
| | Disposition time civil and commercial litigious cases | ** | *** | *** | | *** | *** | *** | | | | | | | |
| | Disposition time civil, commercial and administrative cases | | | *** | | | *** | *** | *** | | | | | | |
| | Time for enforcing contracts | | *** | *** | | *** | *** | *** | *** | *** | | | | | |
| Quality | Crime is effectively controlled | *** | *** | * | | *** | ** | *** | *** | *** | *** | | | | |
| | People do not resort to violence | *** | *** | | | *** | *** | *** | *** | *** | *** | *** | | | |
| | Civil justice is free of improper government influence | *** | *** | *** | ** | ** | *** | | * | | *** | *** | *** | | |
| | Civil justice is effectively enforced | *** | *** | | * | *** | *** | *** | *** | *** | *** | *** | *** | *** | |
| | Criminal adjudication system is timely and effective | *** | *** | ** | ** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** |

r= >0.8 >0.6 >0.4 >0.2 >0.05

Note: Significance levels: * p<0.05, ** p<0.01, *** p<0.001. The correlation coefficient (r) is indicated in absolute terms (modulus).

Figure A.3 shows the observed Eigenvalues[20] of all the factors emerging from this analysis (represented by a dashed line)[21]. In order to determine how many factors to retain, the observed Eigenvalues were compared against those of a randomly generated dataset (represented by a dotted line in the graph)[22]. Only those factors with an Eigenvalue above the dotted line were retained.

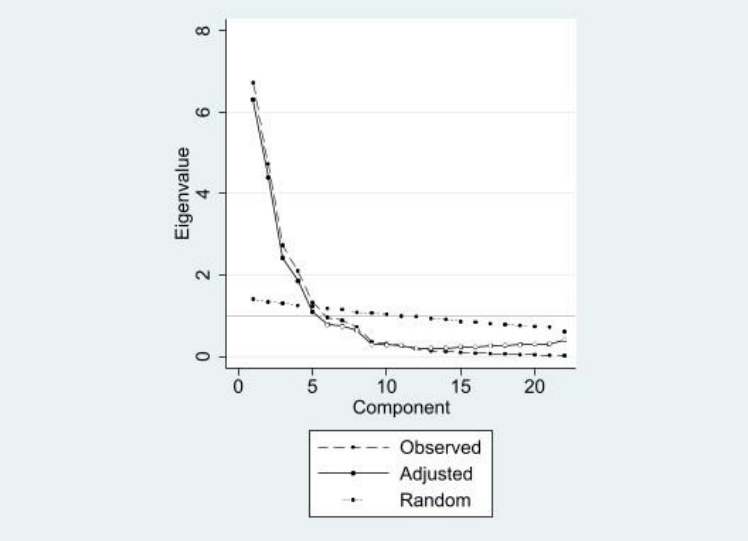
Following Figure A.3, five factors emerge, explaining 80% of the total variance of the original data, with factor 1 alone accounting for 30%. Twenty-two of the 67 indicators were included in the results, since a large proportion were discarded due to high correlation with others from the same service or due to having a small number of observations. This resulted in a total 156 observations for 13 OECD countries (Australia, Austria, the Czech Republic, Denmark, Finland, Hungary, Italy, Japan, the Netherlands, Norway, Slovenia Spain, and Sweden). A larger proportion of indicators of "quality" than of responsiveness and access were included in the results. Similarly, a larger number of health indicators were retained compared with justice and education indicators.

[20] Eigenvalue is a measure of the total amount of common variance between variables explained by a single factor.

[21] Note that the results of the factor analysis would likely differ if more observations (rather than imputed data) and/or different indicators were included.

[22] The random dataset was generated using the "paran" package from Stata (Dinno, 2001[47])

## Figure A.4. Scree plot




Note: The dashed line shows the observed Eigenvalues, the straight line shows the adjusted Eigenvalues and the dotted line shows the randomly-generated Eigenvalues.

Table A.4 shows that the indicators' relative contribution (sign) to the factors are in line with expectations in the majority of cases; however, the score of some factors may be inverted. For example, 30-day mortality rate following acute myocardial infarction (AMI) hospitalisation has a positive (+) sign instead of a negative (-) sign as expected. Moreover, health care coverage has a negative (-) sign instead of a positive (+) sign as expected in factor 1. This indicates that factor 1 is inverted (a positive value in this factor entails lower satisfaction), this is also the case for factor 5, where the percentage of NEET has a positive sign and "people can access and afford civil justice" has a negative one.

Only two indicators out of the 22 have the opposite than expected sign in the factor in which they have their highest loadings: number of practising physicians (which has a positive sign on access, similar to out-of-pocket medical expenditures, although these indicators were expected to have opposite contributions), and proportion of time principals spend talking to parents (which has a positive sign, while "people can access and afford civil justice" and "people can access and afford civil justice" have a negative sign in factor 5, yet the three of them were expected to have the same sign).

## Table A.1. Factor analysis results

| Service | GAAG Dimension | Expected sign | Indicator | Factor1 | Factor2 | Factor3 | Factor4 | Factor5 | Uniqueness |
|---|---|---|---|---|---|---|---|---|---|
| Health | Quality | - | 30 day mortality rate following AMI hospitalisation | 0.80 | 0.07 | -0.34 | 0.21 | 0.14 | 0.25 |
| Education | Quality | - | Percentage variation in reading performance explained by socioeconomic background | 0.74 | 0.01 | 0.22 | -0.05 | -0.08 | 0.33 |
| Health | Quality | - | Mortality rate cerebrovascular disease (stroke) | 0.71 | 0.14 | -0.03 | -0.34 | -0.30 | 0.15 |
| Health | Quality | - | Asthma hospital admissions per 100k population | 0.65 | 0.17 | -0.37 | 0.34 | -0.09 | 0.32 |
| Justice | Quality | + | Civil justice is free of improper government influence | -0.80 | 0.20 | 0.17 | 0.41 | -0.15 | 0.11 |
| Health | Access | + | Health care coverage | -1.08 | 0.28 | -0.13 | -0.21 | 0.16 | 0.06 |
| Health | Quality | - | CHF hospital admissions per 100k population | 0.49 | 0.53 | 0.25 | -0.20 | -0.02 | 0.06 |
| Education | Access | + | Enrolment rate at ages 3 and 4 | 0.17 | -0.97 | 0.33 | 0.02 | -0.13 | 0.10 |
| Health | Quality | - | Mortality rate of acute myocardial infarction (heart | 0.13 | 0.60 | 0.38 | 0.04 | -0.34 | 0.21 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | attack) | | | | | | | |
| Health | Quality | - | Diabetes hospital admissions per 100k population | -0.07 | 0.75 | -0.09 | 0.34 | 0.00 | 0.33 |
| Health | Access | - | Out-of-pocket medical expenditures | 0.45 | 0.08 | 0.55 | 0.45 | 0.43 | 0.13 |
| Education | Quality | - | Share of students below level 2 proficiency in science | 0.22 | -0.30 | 0.64 | 0.01 | 0.32 | 0.25 |
| Education | Quality | + | Mean score in science | -0.16 | 0.31 | -0.83 | 0.11 | -0.06 | 0.10 |
| Education | Access | + | First time tertiary entry rate | -0.20 | -0.14 | -0.69 | 0.16 | 0.13 | 0.29 |
| Health | Access | + | Practising physicians per 1k population | -0.46 | 0.10 | 0.90 | 0.27 | 0.09 | 0.13 |
| Health | Quality | - | 30 day mortality rate following ischemic stroke hospitalisation | 0.21 | 0.31 | -0.08 | -0.61 | -0.11 | 0.37 |
| Justice | Quality | + | People do not resort to violence to redress personal grievances | 0.13 | 0.18 | 0.02 | 0.72 | -0.48 | 0.16 |
| Justice | Responsiveness | - | Enforcing Contracts - Time (World Bank Doing Business Indicators) | -0.05 | -0.04 | -0.06 | -0.95 | 0.09 | 0.12 |
| Education | Responsiveness | - | Percentage of young people (aged 15-29) years not in education, employment or training | 0.13 | -0.03 | 0.09 | -0.33 | 0.78 | 0.15 |
| Education | Responsiveness | + | Proportion of time lower secondary school principals spend talking to parents | -0.25 | 0.06 | 0.12 | -0.04 | 0.96 | 0.11 |
| Justice | Access | + | Alternative dispute resolution mechanisms are accessible, impartial, and effective | -0.33 | -0.31 | -0.22 | 0.35 | -0.35 | 0.30 |
| Justice | Access | + | People can access and afford civil justice | -0.39 | -0.38 | 0.22 | 0.03 | -0.48 | 0.35 |
| | | | | | | | | | |
| | | | Eigenvalue | 6.73 | 4.73 | 2.73 | 2.10 | 1.33 | |
| | | | Cumulative variance explained | 0.31 | 0.52 | 0.65 | 0.74 | 0.80 | |

Note: the values in columns are the rotated factor loadings. The highlight colours indicate the grouping of indicators into factors following their highest loading (green = factor 1, yellow = factor 2, magenta = factor 3, light blue = factor 4, and grey = factor 5).

For each of the five factors, a score for each observation was computed by means of regression, in order to analyse the correlation with satisfaction with services. Satisfaction with services is the average of satisfaction with health care, with education and confidence in justice, from Gallup World Poll. Factor 1 has a correlation of r=-0.55 with satisfaction with all services combined, factor 4 has a correlation of r=0.62 and factor 5 has a correlation of r=-0.58, all of them significant at the 1% level (Table A.5). Interestingly, no indicator from health has its highest loading on Factor 5, yet the correlation with satisfaction with health is almost as strong as the correlation with satisfaction with all services combined. Factors 2 and 3 do not have a correlation with satisfaction with services. By comparison, government expenditures as a share of GDP has a small correlation (of r=0.20) with satisfaction with all services combined, significant at the 1% level.

## Table A.2. Correlation between factor scores and satisfaction with services

| Variables | Factor 1 | Factor 2 | Factor 3 | Factor 4 | Factor 5 | Government expenditures as a % of GDP |
|---|---|---|---|---|---|---|
| Satisfaction with all services combined | -0.55*** | -0.04 | -0.05 | 0.62*** | -0.58*** | 0.20** |
| Satisfaction with education | -0.46*** | 0.11 | -0.04 | 0.16* | -0.65*** | 0.27*** |
| Satisfaction with health | -0.55*** | -0.11 | -0.06 | 0.38*** | -0.53*** | 0.04 |
| Confidence in justice | -0.41*** | -0.07 | -0.03 | 0.80*** | -0.39*** | 0.19** |

Note: Satisfaction with services is the average of Gallup World Poll's satisfaction with quality of health care, satisfaction with the education system and the schools and confidence in the judicial system and courts. The values are Pearson's correlation coefficients, the correlations are computed only for observations (country/year) that were retained in the factor analysis. Significance levels: * p<0.05, ** p<0.01, *** p<0.001

Table A.6 shows the results of a regression between the scores for each factor. All coefficients are significant (at the 1% or 5% level), except for factor 4. As expected, the coefficients for factor 1 and factor

5 are negative. The results of this regression were used to predict satisfaction scores for each country and year, as shown in Figure 1.

### Table A.3. Regression between satisfaction with services and factor scores

| Satisfaction with all services combined | Coef. | St.Err. | t-value | p-value | [95% Conf | Interval] | Sig |
|---|---|---|---|---|---|---|---|
| Factor 1 | -7.075 | .389 | -18.21 | 0 | -7.843 | -6.308 | *** |
| Factor 2 | 3.695 | .402 | 9.19 | 0 | 2.9 | 4.489 | *** |
| Factor 3 | 5.84 | .379 | 15.39 | 0 | 5.09 | 6.59 | *** |
| Factor 4 | -.57 | .454 | -1.26 | .211 | -1.466 | .326 | |
| Factor 5 | -.948 | .431 | -2.20 | .029 | -1.799 | -.097 | ** |
| Constant | 67.651 | .407 | 166.08 | 0 | 66.846 | 68.456 | *** |
| | | | | | | | |
| Mean dependent var | 67.69 | | SD dependent var | 11.14 | | | |
| R-squared | 0.81 | | Number of obs | 153 | | | |
| F-test | 190.49 | | Prob > F | 0.00 | | | |
| Akaike crit. (AIC) | 930.50 | | Bayesian crit. (BIC) | 948.67 | | | |

Note: *** p<.01, ** p<.05, * p<.1