Tilburg University

Weighting justice

Gramatikov, M.A.; Laxminarayan, M.S.

Link to publication in Tilburg University Research Portal

Weighting Justice: Constructing an Index of Access to Justice

Martin Gramatikov

Tilburg University
Tilburg Institute for Interdisciplinary Studies of Civil Law and
Conflict Resolution Systems (TISCO, Hague Institute for the Internationalisation
of Law (HIIL)
m.a.gramatikov@uvt.nl

&

Malini Laxminarayan

Tilburg University
Tilburg Institute for Interdisciplinary Studies of Civil Law and
Conflict Resolution Systems (TISCO), Hague Institute for the Internationalisation
of Law (HIIL)
M.S.Laxminarayan@uvt.nl

Abstract

The methods and procedures for measuring the performance of a legal system are debatable issues. Legal indexes, although not common, do exist as a means for measurement and evaluation of various legal phenomena. This paper outlines one method using the Access to Justice Index, which is a procedure for aggregating the relevant data on the costs, the quality of the procedure and the quality of the outcome of a given legal path. This data is a result of the various justice theories that measure experiences of the user when attempting to solve a dispute. The paper investigates in detail the need for and applicability of weights in order to adjust the observed data. After assessing the possible methods of extracting and applying weights, we conclude that factor analysis is in fact the most practical and reliable technique. The option of non-weighting is also briefly considered, and sheds light on the possibility for leaving the data in its original form. The paper concludes with a discussion on the validation of the index and a case study of consumer disputes in the Netherlands to illustrate its performance.

"If you cannot measure it, you cannot improve it" Lord Kelvin

I.   Legal indexes

Different factors fuel the growing need for the measurement of legal phenomena. Globalization and worldwide competition, universal character of human rights, 'judicial borrowing', and donors' investments in legal reforms are only a few of these factors. How to measure the performance of a legal system is a contentious issue. Looking at the inputs of the system could provide some knowledge on the general legal infrastructure, but the input-based approach could, at best, provide an approximation for the performance levels. The outcomes of the legal system are a more valid representation of its ability to solve problems, provide legal certainty and reinforce the social order. Again, the question remains what set of indicators better gauges the outcomes of a legal system. A popular approach to assess the outcome is to concentrate on the Rule of Law (RoL) paradigm. The RoL, however, is a complex and unobservable part of the social world. Its existence and strength could be judged only after an assessment of numerous indicators, including independence of the judiciary, equal and consistent application of the law to similar cases, and accountability of the government.

Accessibility of the national justice systems or accessibility of a single dispute resolution procedure is a particularly powerful indicator of the fairness, effectiveness, and inclusiveness of the legal systems. Access to justice, however, is also a complex phenomenon consisting of different aspects and meanings. These phenomena could vary across time, jurisdiction, social class or legal culture. One plausible strategy to measure such a rich and diverse domain is to use a set of standard indicators. Each individual indicator would measure separate parts of the general concept. In order to cast light on the whole, rather than on parts, the indicators should be combined into an index.

Every index has two general purposes – integration and parsimony[1]. Integration requires the sum of the parts to provide more knowledge than individual parts. For instance, the World Bank's Worldwide Governance Indicators (WGI) combines a vast amount of individual indicators together to assess the level of governance in the national states[2]. In real life, however, governance is only a vague concept which cannot be observed and measured directly. The theory of the indicators states that it is an observable phenomenon, which could be assessed by looking at observable sub-indicators or markers. The WGI distinguish six dimensions of the governance, one of which is the RoL. It is a complex and unobservable phenomenon defined as "....the extent to which agents have confidence in and abide by the rules of society, and in

---

[1] ELIZABETHANN O'SULLIVAN & GARY R. RASSEL, Research Methods for Public Administrators (Longman Third Edition ed. 1999); LEA BISHOP SHAVER, *Defining and Measuring A2K: A Blueprint for an Index of Access to Knowledge*, 4 I/S: A Journal of Law and Policy for the Information Society 235, (2008).
[2] DANIEL KAUFMANN, et al., *Governance Matters VII: Aggregate and Individual Governance Indicators, 1996-2007 available at* http://ssrn.com/paper=1148386.

particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence"[3].

Apparently, the enforcement of contracts and the likelihood of crime measure different things on different scales. However, the WGI team believes that these are aspects of the higher-level RoL concept, and, thus, treats their scores as parts of the composite index. Therefore, the purpose of the index is to integrate the different parts of the whole into a single value. Dealing with one representation of the idea of RoL instead of with its dimensions provides parsimony – clarity and a simplified explanation of the observations.

Similarly, Freedom House collects data and computes its Freedom in the World Index[4]. RoL is considered an important dimension in the broader idea of *freedom.* As in the WGI, the Rule of Law is treated as a complex phenomenon which can only be observed through proxy indicators. Independency of the judiciary, civic control over police, the extent of legal protection against external intrusions, and fairness of the law are the four ingredients of the RoL, according to the theoretical framework of Freedom House. Each of these ingredients is put into operation through a set of items in which data is collected. In order to integrate the fragments of the RoL concept, researchers combine the items into higher-order dimensions. The dimensions are then aggregated into a composite index which provides information on the level of the RoL in a particular country.

Several other global or regional indexes measuring legal constructs are related to access to justice[5]. The Heritage Foundation measures economic freedom. One of the components of this index is the property rights defined as an "an assessment of the ability of individuals to accumulate private property, secured by clear laws that are fully enforced by the state"[6] . The American Bar Association World Justice Project develop and test an index, which aims to measure the countries' adherence to the RoL[7]. Four broad factors (called bands) are contemplated as proxies of the index, with the third band closely connected to the accessibility of the justice system. Its content is described as "...the accessibility, fairness and efficiency of the process by which the laws are enacted, administered and enforced"[8].

In this paper, we discuss the substance and the procedure for computation of the Access to Justice Index. Both aspects are of critical importance for the usefulness of the index. A rigorously computed index of data which does not properly represent the measured idea may end up being a worthless use of time and resources. The opposite is also true – even if the data validly measure components of access to justice, a

---

[3] Id. at *7.*

[4] FREEDOM HOUSE, Freedom in the World 2007: The Annual Survey of Political Rights and Civil Liberties   (Rowman & Littlefield Publishers. 2007).

[5] See more on global and regional indexes that measure legal phenomena can be found in CHRISTIANE ARNDT & CHARLES OMAN, Uses and abuses of governance indicators (Development Centre of the Organisation for Economic Co-operation and Development. 2006).

[6] KIM R. HOLMES, et al., 2008 Index of Economic Freedom 41 (The Heritage Foundation 2008).

[7] See http://www.abanet.org/wjp/

[8] THE WORLD JUSTICE PROJECT, *Rule of Law Principles and World Justice Project Overview.* (2008), at http://www.abanet.org/wjp/docs/abawjpresourcekit.pdf.

flawed index construction will cast doubts on the outcome. In order to mitigate these challenges, we will first review the substantial part of the index. All three indicators will briefly be introduced and explained. Then, alternative options for the aggregation of these parts into one composite index will be investigated. The options will be tested on a data set from a survey of the path to justice in the Netherlands.

In chapter 1, the concept of an index and the content and formulation of some of the most popular legal indexes were discussed. The remainder of the paper proceeds as follows: The second part of the paper examines the decisions that have to be made in order to aggregate the pieces of data into one index. Special emphasis is given to the issue of weighing the data. In addition, three alternative strategies to derive weights for the quality of the procedure and the quality of outcome will be discussed. The options will then be applied to the observed data in order to study the impact of the weights. Finally, we discuss a strategy for assessing the validity of the composite index.

II. Measuring access to justice

In the last decades, a growing number of studies have applied theories and methods from humanities, social sciences and economics to understand different legal phenomena. Access to justice is one of the legal domains that gathered significant inter-disciplinary attention. Increasingly, the accessibility of the legal systems, and in particular the dispute resolution systems, are measured through quantitative methods. Since the 1970s, researchers employed empirical strategies to estimate the legal needs that exist in society[9]. In these early studies, legal needs were expected to fall under the category of professional legal advice. Although from different viewpoints, access to justice is quantitatively examined in studies measuring the costs and delay of litigation procedures[10].

Access to justice has been measured in varying ways, often dependant on the discipline carrying out the research. Psychologists focus on the procedural and distributive justice needs of users[11], while economists are more interested in the costs associated with a given procedure. Victimologists are concerned with the negative repercussions crime victims suffer while gaining access to justice[12]. For economists, social scientists and lawyers alike, there is a desire to link access to justice to those

---

[9] AMERICAN BAR ASSOCIATION, Comprehensive Legal Needs Study. (1994); BARBARA A. CURRAN, The Legal Needs of the Public: The Final Report of A National Survey (American Bar Foundation. 1977).

[10] HERBERT M. KRITZER, et al., *Understanding the Costs of Litigation: The Case of the Hourly Fee Lawyer*, Summer 1984 American Bar Foundation Research Journal 559, (1984); DAVID M. TRUBEK, et al., *The Costs of Ordinary Litigation*, 31 UCLA L. Rev. 72, (1983); EDGAR LIND, et al., The Perception of Justice. Tort Litigants' Views of Trial, Court-Annexed Arbitration, and Judicial Settlement Conferences (1989).

[11] S. ALEXANDER & M. RUDERMAN, *The role of procedural and distributive justice in organizational behavior*, 1 Social Justice Research, (1987); LAURA KLAMING & IVO GIESEN, *Access to Justice: The Quality of the Procedure available at* http://ssrn.com/paper=1091105; JOHN THIBAUT & LAURENS WALKER, Procedural justice: A psychological analysis (Erlbaum. 1975).

[12] ULI ORTH, *Secondary Victimization of Crime Victims by Criminal Proceedings*, 15 Social Justice Research, (2002); JOANNA SHAPLAND, et al., Victims in the criminal justice system (Gower. 1985).

suffering from social inequality. Other topics of interest include informal justice methods, access to legal aid and public interest law[13].

In the 1990s, research on access to justice shifted towards the direction of broader inquiry on the sociological aspects of access to justice. In an influential study, Professor Hazel Genn surveyed the response strategies to non-trivial problems that might have a resolution with legal means[14]. The so-called "justiciable events methodology" gained popularity and has been replicated over time and jurisdictions[15]. Its power is based on a comprehensive assessment of the impact that legal problems have in everyday life. What the "justiciable" research methodology still cannot deliver, however, is a comparison of the results across jurisdictions. Although the inferences are based on perception-based data, the various research projects split into different categories, including types of problems, period of interest and data collection method.

A different approach to the quantitative assessment of access to justice is proposed by the "Measuring Access to Justice" (MA2J) project[16]. The project develops a research methodology which measures access to justice through the perceptions and attitudes of people who have travelled a path to justice. As compared with the justiciable events studies, the MA2J diverges significantly – its units of measurement are individuals who had a legal problem and acted to solve it with the means of state or non-state intervention. In fact, three major aspects of the experiences of the users are measured – the costs of justice[17], the quality of the procedure[18] and the quality of the outcome[19]. Each of these three pillars of the user's experiences is modelled as a multi-faceted indicator consisting of sub-indicators. The individual scores of the sub-indicators form the scores of the three higher level indicators. At the next level, the project intends to aggregate the information on costs, quality of the procedure and quality of the outcome into one composite figure. This single value, the Access to Justice Index, should provide focused information about the measured paths to justice.

A. Usefulness of the index

First, we should clearly state the reasons for the index and its usefulness in examining procedures individually and in comparison to another. Individual evaluation may offer

---

[13] MAURITS BARENDRECHT, et al., *How to Measure the Price and Quality of Access to Justice? available at* http://ssrn.com/paper=949209.

[14] HAZEL GENN, Paths to Justice. What people do and think about going to law? (Hart Publishing. 1999).

[15] AB CURRIE, The Legal Problems of Everyday Life (2007);HAZEL GENN & ALAN PATERSON, Paths to Justice Scotland. What People in Scotland Do and Think About Going to Law (Hart Publishing. 2001);PASCOE PLEASENCE, et al., Causes of Action: Civil Law and Social Justice (Stationery Office. 2004);MARTIN GRAMATIKOV, Multiple Justiciable Events in Bulgaria (Legal Services Research Centre ed., 2008);BEN C. J. VAN VELTHOVEN & MARIJKE TER VOERT, Paths to Justice in the Netherlands. Looking for signs of social exclusion. 28 (2004).

[16] BARENDRECHT, et al., *supra note 13*; MARTIN GRAMATIKOV, et al., Measuring the Costs and Quality of Paths to Justice: Contours of a Methodology (SSRN 2008).

[17] MARTIN GRAMATIKOV, *A Framework for Measuring the Costs of Access to Justice*, 2 J. Juris. 111, (2009).

[18] KLAMING & GIESEN, *supra note 11*.

[19] JIN HO VERDONSCHOT, et al., *Measuring Access to Justice: The Quality of Outcomes available at* http://ssrn.com/paper=1298917.

less information than comparative values, but insights can still arise concerning the need for improvement. Furthermore, the aim of the index is – once data has been collected and analyzed – to compile rankings of procedures in a relatable manner. Understanding what is relatable or comparable (between jurisdictions or between procedures) will be elaborated on below. Using these evaluations and comparisons can lead to policy implications; the index can serve as a learning tool among jurisdictions.

Evaluation

The idea behind indexes is simple: to summarize several indicators into one numerical score. One objective of the index is for the evaluation of a given procedure. If a neutral is concerned with user perceptions of his or her procedure, the index can provide numerous insights into the various components of the process. Furthermore, if a user is curious about a given procedure, and has access to this information, he or she may obtain a better understanding of what lies ahead. Although costs may be higher in one case, a more favourable outcome may outweigh this shortcoming. Each component can be compared with the use of the sub-indexes. In addition, this evaluation can be used for comparisons and policy evaluations, as we will discuss.

Comparison

Both comparisons among countries or paths and comparisons over time can offer various policy implications. Comparison over time can illustrate whether or not there have been improvements in certain procedures. For example, in the field of victimology, legal provisions have recently taken place and will continue to do so at both national and supranational levels. The effects of these changes can be measured by victim satisfaction and their perceptions of the quality of the experiences, and, consequently, the rankings conferred to various procedures. Legislation reform can be established in one jurisdiction; actual implementation, however, may paint a much different picture. This notion can be illustrated by one example in the Dutch criminal justice system. In the past decade, various guidelines (as stated by the 1995 Terwee Act)[20] have been established which improve information (informational justice items), participation (procedural justice items), and treatment of victims[21] (interpersonal justice items). In this instance, not having only one composite index, but instead several sub-indexes for each type of justice, will prove to be an advantage. Furthermore, longitudinal studies using these indexes would provide for valuable insights into the policies which have been established. More attention can be devoted to those areas which may be lacking acceptable values. Understandably, the index will not be valuable unless legal problems are either similar between different paths in the same country or similar legal problems between countries follow the same paths. For example, labour issues in the Netherlands will not be compared to consumer disputes in Belgium. Rather, the aim is to extract weights (discussed in detail in the latter part

---

[20] Wet Terwee, Victim Act Terwee, Gazette 1995, no. 65., incorporated in the Dutch Code of Criminal Procedure of 15 January 1921, Stb. 14, amended 6 December 1995, Stb. 592

[21] MARION ELEONORA INGEBORG BRIENEN & ERNESTINE HENRIËTTE HOEGEN, Victims of crime in 22 European criminal justice systems : the implementation of Recommendation (85) 11 of the Council of Europe on the Position of the Victim in the Framework of Criminal Law and Procedure : proefschrift (Wolf Legal Productions (WLP)/Vidya in cooperation with the Global Law Association. 2000).

of this paper) for each path to justice, and, consequently, draw conclusions from those paths which are in fact comparable. Comparable data has further implications for policy changes, suggesting that countries have the ability to look to each other's legal systems and procedures for administering justice.

B. Access to justice indicators

Our theory of the Access to Justice Index is based on the belief that the users' perspective is appropriate ground for assessing the extent of accessibility. A focus on the experiences of the users of justice means that the units of analysis are particular paths to justice. All conclusions and inferences will be applicable to the measured path to justice. Our definition of a path to justice is "a commonly applied process which users of justice address in order to cope with their justice needs"[22]. Both formal and informal processes could be positioned under the broad scope of the definition of a path to justice. A path begins when the user first addresses the process. This first active and deliberative involvement of the user into the process could be explicated through different acts – searching for information, acceptance of advice, filing documents, etc. The path ends when the user receives an outcome from the process. An outcome could be a final decision by a mediator, a joint agreement of the parties, or an end to the process because one of the parties quits.

The experiences of the users of justice are measured against three indicators: the costs of the procedure, the quality of the procedure, and the quality of the outcome. Each indicator is built by a complex structure of sub-indicators. For instance, the costs of the procedure are defined as the resources, which the user would need to travel from the beginning to the end of a path to justice. Within this indicator, a set of sub-indicators reflect different types of procedural costs: out-of-pocket expenses, time, and other opportunity and intangible costs (stress, emotions, etc.). Similarly, the two quality indicators consist of lower level sub-indicators that measure its specific facets.

---

[22] MARTIN GRAMATIKOV, *Methodological Challenges in Measuring Cost and Quality of Access to Justice available at* http://ssrn.com/paper=1099392.

Figure 1: Access to Justice Index
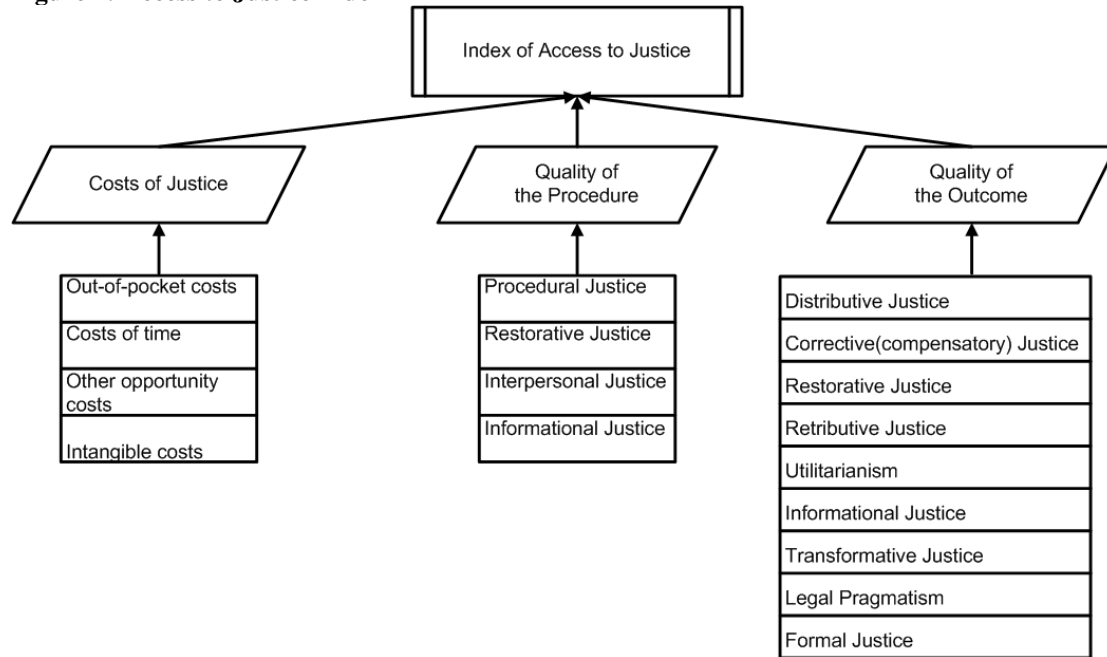
Figure 1 outlines the components of the Access to Justice Index. In the table below we outline the sub-indicators of access to justice in more detail. An elaboration of these concepts can be found in relevant papers[23]. Important to note, however, is that not all indicators are universal across users. To give an example, while indicators of utilitarianism may be applicable to users of the criminal justice system, the same may not hold true for other legal problems.

Table 1: Indicators of the Costs and Quality of Paths to Justice

Costs of the Procedure		
Indicator	*Description*	*Examples*
Out-of-pocket expenses	The monetary amount spent on transactions during and as a result of the proceedings	Lawyer fees, expert fees, filing fees, transportation fees, bailiff and witness fees, notary fees, costs for communication
Time	Time spent dealing with the procedure	Searching for a legal advisor, collecting information, contacting professionals, travelling, awaiting/attending hearings, waiting in queues
Other lost opportunities	The cost of lost opportunities due to the proceedings and their possible lengthiness	Lost income, devaluated resources, losing a job opportunity
Intangible costs	On their paths to justice, people tend to expend emotions, suffer stress,	Stress, negative emotions such as frustration, fear, disappointment or anger, loss of relationships

---

[23] KLAMING & GIESEN, *supra note 11*; GRAMATIKOV, *supra note 17*; VERDONSCHOT, et al., *supra note 19*; MALINI LAXMINARAYAN, Measuring Crime Victim's Paths to Justice: Developing Indicators for Costs and Quality of Access to Justice (Tilburg University 2008).

	become depressed or experience deterioration in their relationships with significant others	
Quality of the procedure		
Indicator	*Description*	*Sub-indicators*
Procedural Justice	Fairness perceptions of users regarding the processes that are utilized to resolve disputes and allocate resources	Process control, decision control, consistency, bias suppression, accuracy, ability to correct, ethicality
Restorative Justice	Concerned with the harm that has been caused by the legal problem and attempts to offer reparation to the user of justice	Opportunity to ask the other party for an explanation and recognition
Interpersonal Justice	The extent to which people are treated with politeness, respect, and propriety	Politeness, respect, propriety, respect for rights
Informational Justice	The validity of information provided by decision makers as the foundation of the decision making process	Honesty, explanation of rights and options, as well as whether the explanation was timely, understandable, and in need of clarification
Quality of the Outcome		
Indicator	*Description*	*Sub-Indicators*
Distributive Justice	The allocation of resources among individuals with competing needs or claims	Equity, equality, need
Corrective (compensatory) Justice	When one person is wrongfully injured by another, the injurer must make the harmed party whole	Compensation
Restorative Justice	Looks to the future and the best means to achieve reparation of harm, including elements of apology, shame and repair of relationships	Opportunity given to the offender to show remorse and to accept responsibility, the degree of reparation of emotional and monetary harms, closure, alleviation of fear
Retributive Justice	An infliction of proportionate loss and pain to the injurer is necessary to achieve justice	Just deserts
Utilitarianism	Social harmony can be attained via the prevention of future harm	Deterrence and incapacitation
Informational Justice	The validity of information provided by decision makers	Outcome justification

	as the foundation of the decision making process	
Transformative Justice	The future is a main concern, in addition to transformations with the self, self-in-relationships, and self-in-society	Considering parties' interests and re-building relationships
Legal pragmatism	Concerned with facts and consequences	Pragmatic outcome, consequences taken into account
Formal Justice	Similar cases are defined by the situation, not the person – i.e., treat similar cases alike	Similar outcomes and ability to compare outcomes with related others

III. Weighting the index items

One way to compute the Access to Justice Index is to aggregate the values of all items, thus measuring a higher-order concept. The outcome means will reflect the score of the unobservable variables. At the next level, the same computation could be performed in order to discover the scores of the three basic indicators (costs, quality of the procedure, and quality of the outcome). In a third round of aggregation, the composite index of the accessibility of the measured path to justice could be estimated through the mean of the three main indicators. Such a routine uses the non-weighted original values of the items. It assumes that the data reliably measures the indicators allowing the studied sample to formulate reliable inferences.

Alternatively, the two assumptions discussed above may be relaxed. They could either be uncertain that the instrument items measure the concepts of interest well, or the sample deviates systemically from the population. In both cases, weights could be used for adjusting the original values. In essence, the weighting is a procedure of assigning different degrees of importance to the original data depending on additional information. In this chapter, we review the possible alternatives for deriving and applying weights to the items and indicators of the Access to Justice Index.

A. *Sample weights*

The weights could reflect three distinct aspects of the reliability of the index components. Most often, weights are used to adjust the sample toward some known characteristics of the population of interest. In order to adjust the sample data, underrepresented groups are weighted more heavily in order to influence the data in proportion to their participation in the population. The focus of this paper, however, is on item weights and how this procedure can be achieved.

B. Item weights

Weighting items is more challenging, as it attempts to collectively measure an unobservable indicator. In the case of the sample adjustment, the weights are applied to decrease the differences between the sample and the population from which it has been drawn. Weights could also be applied to the items of the research instrument. The logic of this type of weighting lies in the unknown importance of the items. Our theoretical framework determines what items are likely to measure an indicator, but it does not indicate how well each item performs. When we measure the costs and quality of a path to justice, we hypothesize that the users use the criteria set out in Figure 1. However, we do not know whether the distributive justice impacts the perception of the procedure quality more than the informational justice. It is an empirical question to ascertain the magnitude of importance of these criteria.

The assumption that items and indicators could vary in their importance calls for an estimation of the parameters of such importance. But what exactly does importance of an item for measuring justice indicators mean? Two principally different answers could be given to this question. First, subjective importance implies that the particular users of justice have a preference scale on which they estimate the relationships between items. Thus, parties in two different types of legal processes may have different preferences for the criteria used to measure the quality of the outcome. For instance, depending on circumstances and personal preferences, the user could place greater value on the distributive or informational aspects of the outcome. These weights are further referred to as stated preferences weights.

A second method to answering the estimation question takes a radically different approach. Instead of searching for the answer in the users' preference scales, it assumes a more theoretical stance. After the important indicators are into items, this second approach attempts to discover how well the items measure the unobservable indicator. In a sense, it disregards the personal preferences, instead relying on the validity of the measurement tool. For brevity, the weights that reflect the connection with the unobservable indicator will be called extracted association weights.

Both approaches for justifications of item weights will be discussed below. Their advantages and disadvantages could be studied on two levels. One is the conceptual level, where it must be decided which approach is more coherent with the research questions that the Access to Justice Index must answer. The two weighting paradigms reflect different technique of uncovering the importance of the items. One method is based on personal choice; the other is grounded in the ability of the item to measure the latent construct.

Apart from the theoretical differences, the rationale of the two approaches must also be assessed from an empirical perspective. Item weights are not easy to operationalize, measure, and interpret. Often, estimation of item weights requires complicated research designs and procedures. The reliability of the outcomes of the weight extraction is of major concern, and thus should be factored into the discussion of the advantages and disadvantages of the *stated preferences weights* and the *extracted association* weights.

1.   Stated preferences weights

Weights based on the stated preferences regarding the relative importance of an item assume that users place different levels of importance on individual items. For instance, the restorative effect of the outcome could be more significant than its distributive consequences. Or, at another level, the quality of the procedure could be more important than the quality of the outcome. It is possible that the importance function is dependent on the personal characteristics of the respondent. However, it could also be that preferences are influenced by belonging to social groups, types of problems, or types of procedures.

People from different social groups could systematically place dissimilar preferences on individual indicators. Economic position, education or political views could interact with the preferences on the items that measure costs and quality of paths to justice. It is a plausible hypothesis that for those with more resources, the quality of the procedure matters more than that of the outcome. Alternatively, the less fortunate will value the outcome more because of the higher marginal utility. These hypotheses, however, could prove useless because of the many possible interactions between individual characteristics and perceptions on the justice. For instance, people with more resources normally tend to have a better education and know more about their rights. As such, they could place even stronger emphasis on the quality of the outcome.

The importance of the items may also be impacted by the type of process used. State organized paths to justice are more formal, causing users to focus more on the need to receive attention and empathy from the neutral. More legal formalism increases duration and costs of dispute resolution, which could mean that costs will be of utmost importance. On non-state paths, such as negotiations or mediation, the quality of the outcome could be perceived by users as being more important.

2.   Extracted association weights

Many of the indicators for measuring access to justice cannot be observed directly and have to be measured through proxy indicators. While each of the proxy variables captures part of the unobservable variable of interest, they may also reflect other facts and perceptions. Some of the items will measure a larger share of the indicator and, therefore, will be better associated with it. The weight of the items could also be based on their abilities to measure the indicator. For instance, we measure interpersonal justice with four items, and it is highly unlikely that these four items contain an equal amount of information on the unobserved factor. If the quality of the procedure is believed to consist of three components (procedural justice, interpersonal justice, and informational justice), it would be very bold to assume that they participate equally in it. It is entirely possible that one of the items[24] is better associated with the quality of the outcome. Weights, based on the association of the item with the unobservable indicators, should assign more importance on this item and less importance on the others.

---

[24] These are already aggregated items.

Using extracted association weights, the researcher has more influence on the weighting process. The items' weights depend on the relationship (association) of the item and the unobservable indicator. In order to discover the strength and the direction of the association, the index constructor has to extract the weights. As we saw above, in the stated preferences weights, it is the user of justice (the respondent) who directly determines the weights.

3.   Advantages and disadvantages of weighting the items

Preference of one of the options for extraction of weights should be guided by theoretical and practical considerations. It is likely that both weights are applied to different parts of a measurement tool. For instance, in the Access to Justice Index, both weighting schemes could be used to increase the reliability of the composite index. Below is an example of a theoretically driven choice of weights. The theory is grounded in the characteristics and linkages of the index indicators described above.

At the bottom level of the index (Figure 1), where the unobservable criteria are measured by proxy variables, the use of extracted association weights will be more feasible. All proxy-variables are supposed to measure a relatively compact indicator, and it would be difficult for the users of justice to clearly differentiate which item is more important in relation to the others. Application of the extracted association weights relieves the users from the burden of making a difficult choice, and indicates which item is better associated with the unobservable indicator. Thus, the adjustment of the significance of items according to their ability to measure the indicator has sound theoretical fundaments.

An example follows: In the quality of the procedure dimension, the index has four complex indicators. Another eight indicators measure the quality of the outcome (Figure 1). Again, a decision must be made on which weighting scheme to use at this second level of aggregation. The application of extracted association weights is an option, but there is one serious concern. Each of the items is a product of the lower level items. If one wants to find out how the four items measure the quality of the procedure, she must consider the fact that these items inhibit more measurement error due to the aggregation process.

Alternatively, the stated preference weights could be used to improve the reliability at the second level of the index. This means that users should indicate the importance of each of the aggregated items against each other. An item in which the respondents are asked to rate its importance could provide a meaningful preference weight. The disadvantage of this approach is the need for clear operationalization of the aggregated items. Asking respondents about procedural justice, restorative justice, or other highly abstract terms will likely impact the validity of the instrument. Therefore, the use of importance weights could be recommended only when the indicators of interest can be sufficiently explained in a language that will be understood clearly by the respondents.

As stated above, the choice of weights for index items should also be guided by practical considerations. Computing the weights of items requires additional resources in terms of increased samples and larger questionnaires. The stated preferences

14

weights seem to be significantly more challenging. Either the respondents have to reveal their preferences together with the items, or the preferences should be obtained from a different source. While the former approach is appealing, it could undermine the reliability of the measurement. A research instrument which asks for the importance of each item will double its size, inevitably leading to lower response rates and other negative effects. Even more challenging is the extraction of weights from different sources. Later, we discuss the pros and cons of estimating weights in one sample and applying them to the item scores of another.

An interesting related question is whether universal weights could exist for paths to justice. A theory of universal weights must build on the assumption that users of particular paths to justice have stable preferences on the evaluation criteria. Apparently, this assumption should be limited at different levels, such as type of problem, characteristics of the parties, specifics of the particular jurisdiction, the related legal culture, etc.

Unlike the stated preferences weights, the extracted association weights are less demanding in terms of research resources. Normally, the observed data could be easily analyzed with statistical techniques, such as factor analysis or multivariate regression. Below, we give three examples for extraction and application of weights.

IV. Three examples of weighting items

A. Stated preferences weights

In a quasi-experiment, we obtained the stated preferences of users of justice on the importance of the quality of the procedure and the quality of the outcome indicators[25]. A direct result of the measurement is that it provides a hierarchical order, which reflects the subjective importance of the individual items. An item which is ranked as the most significant (e.g. voice) within a group of items (e.g. procedural justice) is deemed as more important than the other items from the same group. Therefore, the rankings can produce meaningful weights, which can be applied to the observed data. Items could also be equally important (e.g. voice and bias suppression). The ultimate outcome from the weighting is that items that were deemed as more important by users will receive higher importance in the composite index of the costs and the quality of access to justice.

1. Challenges

There are several serious disadvantages in the use of stated preferences to weight items of the quality of the procedure and the quality of the outcome indicators[26]. Like the item scores, the ranking of the items have to be estimated empirically. One option is to calculate the items and their relative importance from the same sample of users of justice, meaning that the same questionnaire will collect two types of data.

---

[25] For details on the research design and the methodology of data collection see Appendix 1
[26] See *supra* III.B.3

Alternatively, the item scores could be estimated from one sample of users, while the importance is estimated from a different sample.

Deriving stated preferences from the same or different sample inevitably increases the cost of the indexing construction. Furthermore, in the latter option, there is the risk of the two populations differing in some known or unknown manner. The comparability of the paths to justice could also be an issue. Another challenge may be the moment that stated preferences are a valid and reliable reflection of what users of justice find important, as this could change with time. Change in legislation or external factors might influence the perceptions regarding the relative importance of the items of the quality of the procedure and the quality of the outcome indicators.

B. Association weights extracted with factor analysis

In this section, we focus on the parsimony function of factor analysis and its ability to produce meaningful weights of the items that measure the quality of the procedure and the quality of outcome of paths to justice. Factor analysis (FA) is a statistical procedure for data reduction and identification of latent factors in the data. Thompson adds two additional intentions of the FA – evaluation of score validity and development of theory regarding the construct[27].

There are numerous methods for factor extraction which attempt to achieve the goals of FA from different perspectives. For the purpose of this example, we will use the method of principal component analysis (PCA). The advantage of PCA is its ability to analyze the total variance in the correlation matrix to reduce the data structure. Unlike the alternative methods for extraction of common factors, PCA operates with components extracted from the correlation matrix, and not with hypothetical factors[28]. Another advantage of PCA is that it extracts components from the total variance, whereas the other methods use only the common variance between items.

Unlike most of the parametric statistical procedures, FA does not assume normal distribution of the data. Two assumptions underlie FA. First, it is applicable to continuous variables. Most items in the MA2J measurement instrument use a 5 point Likert scale or are re-scaled to a 5 point scale[29], meaning we can apply FA. Second, the variables must be linearly related to each other. This assumption could be checked through investigation of the co-relation matrix or review of the scatter matrix.

The primary goal of PCA is to reduce a set of correlated items into a few components which are more amenable to interpretation. In the context of the MA2J methodology, the components are expected to be indicators and sub-indicators of the costs and quality of the paths to justice, and the items are the individual questions from the measurement instruments. For instance, PCA could indicate whether the five variables designed to measure the procedural justice primarily measure this factor, or if the data captures other facets of the experience with the procedure. After estimating the number of components, PCA reveals the association between each variable and the extracted component. These correlations could be interpreted as the fit between an

---

[27] BRUCE THOMPSON, Exploratory and confirmatory factor analysis : understanding concepts and applications pp. 4-5 (American Psychological Association 1st ed. 2004).
[28] PAUL KLINE, An easy guide to factor analysis (Routledge. 1994).
[29] The costs are measured at ordinal level and then re-scaled again to ordinal variables.

item and a component. Items with high correlations are deemed to be good measures of the component. On the other hand, items with low correlations measure some other factors, which are not shared with the other items.

The measure of the association between the individual items and the measured indicators could be used as weights when interpreting the results from the MA2J studies. Items better related to the particular indicator will be given more weight in the index variable. Later, we discuss the capability of PCA to provide meaningful weights for the MA2J data. In FA, the measure of extracted communalities is directly related to the factor loadings. The communality of a variable reflects the amount of variance, explained by the extracted factors. If there is only one factor identified in the data structure, the communalities express the variance explained by this single factor. There is a direct relationship between the factor loadings and the communalities - the latter are the sum of the squares of the factor loadings. The application of the communalities or the factor loadings as weights will reward the variables, which are closely associated with the indicator.

A practical question is which measure to use as weight for the observed variables – the communalities or the factor loadings. As the former are a product of the latter, both weights will produce similar, but not equal, results. In order to avoid further complications of the interpretation of the weighted variables, we recommend the use of the factor loadings as weights for the observed data.

1. Challenges

The factor loadings can be used as meaningful weights if there is only one factor extracted. There is the possibility that despite the careful alignment of the measurement instrument with the theoretical framework, the variables in fact measure two or more factors. In this case, the communalities will be the product of more than one factor loading. While this will contradict our theoretical framework, it will also challenge the use of the factor loadings as weights. If the data does not fit the theoretical model well, the validity of the measurement instrument has to be carefully scrutinized.

Another practical challenge is the sample size. Although there is no fixed requirement for the size of the sample, there are certain suggestions. The general advice is that larger samples will allow for more valid and reliable weights extraction.

C. Association weights extracted with multiple regression

Multiple regression is the second statistical procedure that will be reviewed for the purpose of deriving association weights. Briefly stated, standard multiple regression allows for the assessment of the relationship between one continuous dependent variable and numerous independent variables. For example, is satisfaction with the criminal justice system related to several independent variables, such as respectful treatment, participation, being informed and/or sentence severity?

Using multiple regression allows for the identification of a set of predictor variables that, when combined, will provide a useful estimate of the score of the participant's criterion variable. It is possible to predict the score of the dependent variable based on

the scores of several variables. Multiple regression accounts for the variance in the observed scores. The regression equation is as follows:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + \varepsilon$$

Where Y is the predicted or explained value on the dependent variable, X represents the independent variables explaining the variance on Y, and B is the coefficient assigned to each independent variables. These coefficients are the weights attributed to each independent variable and are similar to the factor loadings in FA. Finally, A is the intercept[30].

To find out which variables in the model contribute to the prediction of the dependent variable and to what extent, the beta coefficient is used. Since this research is solely weighting the procedure and outcome – elements which do no need standardization as they are on the same scale – conversion is not necessary and the un-standardized coefficient can be utilized. Applied to the current research, the weighted equation for procedural justice would be as follows:

Y(composite procedural) = A + b_1 (process control) + b_2 (decision control) + b_3 (consistency) + b (bias suppression) + b (accuracy) + b (correctability) +b (ethicality)

1. Challenges

An issue with multi collinearity arises if two or more independent variables are highly correlated with one another, and, therefore would basically be measuring the same phenomenon. This study will utilize the value .75 as a threshold for identifying multi collinearity. Different methods are available to ensure that this limitation does not exist.

This analysis will follow the rule that regression with one independent variable requires 30 observations, and each additional independent variable added to the equation will require an additional 10 observations. If measuring procedural justice, which is calculated using 7 independent variables, a minimum of 90 observations would be necessary.

As seen in Figure 2, there will be a largely skewed value for each variable if multiple regression is used. One possibility is to leave the non significant variables out of the equation. The existence of non-significant values is another issue suggesting multiple regression is an inferior method when compared to FA. Because of the distorted weighted values derived by multiple regression, FA may prove to be the superior method in extracting weights for this research. The other option is to assume that variables 1 and 2 are the primary predictor variables affecting the criterion variable of satisfaction with the procedure. As they are the only variables which are statistically significant, the rest may be eliminated. Multiple regression is furthermore an inferior method due to its need to meet the above mentioned assumptions – in addition to other statistical issues such as normality, linearity and independence of errors – which is often a difficult task.

---

[30] BARBARA G. TABACHNICK & LINDA S. FIDELL, Using multivariate statistics (Allyn and Bacon 4th ed. 2001).

D. Using un- weighted data

The possibility of using an un-weighted index should not be overlooked. Often, index systems rely on equal weighting. In this case, the means will simply be utilized in an equation as follows:

Y(composite procedural) = (1)voice + (1)decision control + (1)consistency + (1)bias suppression + (1)accuracy + (1)correctability + (1)ethicality

Comparing weighted and un-weighted indexes will often yield similar results. This outcome leaves one to question whether a weighted index will be necessary to measure the quality of a path to justice. Weighted and un-weighted scores will be highly correlated, as was the case in Likert's classic study measuring attitudes which found a .99 correlation[31]. Additionally, Nunnally argues that it is difficult to defend any given method as superior to simply summing un-weighted ratings[32]. As discussed above, many legal indexes do not apply any weights to the observed data.

Several issues emerge when weighting the observed data. The assumptions required when conducting statistical analyses were discussed above. As is the case of the stated preferences, which was described using the law student experiment, applying one type of weight to a given criteria cannot be generalized to all types of legal problems. Other methods, such as the multiple regression, produce difficult to interpret weights which cannot be applied without additional transformation. It is also possible that the data severely contradicts the theoretical model which could render the weights meaningless. Finally, there is inevitable discrepancy between the weights derived with statistical methods and the actual preferences among the population of users of specific path to justice.

V.   Application of weights

In this section, we apply the extracted weights to the observed data and analyze the impact of the weights. A good weight is expected to draw the observed data closer to the real properties of the studied phenomena. When the units of measurement are perceptions about the quality of the procedures and the outcomes, it is difficult to find precise benchmarks to assess the performance of the weights. In order to study the performance of the three weighting schemes, we are going to compare the weighted data with the observed data.

Following the discussion above, three types of weights were extracted from two data sets (see Table 2). According to our taxonomy, the factor analysis (FA) weights and the multiple regression (MR) weights are extracted association weights. The third weighting model (Stated Preferences – SP) is based on users' stated preferences (see

[31] RENSIS LIKERT, *A method of constructing an attitude scale*, *in* Scaling: A sourcebook for Behavioral Scientists, (G. Maranell ed., 1974).
[32] JUM C. NUNNALLY, Psychometric theory (McGraw-Hill 2d ed. 1978).

Appendix A for details on the computation of the indexes). Both association weights were extracted from a sample of users of justice who attempted to solve a consumption dispute. Data from a quasi-experiment with law students was used to compute the stated preferences weights.

Despite the fact that the weights were taken from different samples and different legal problems, there is a significant positive correlation between the FA and the SP[33] weights. Positive correlation as high as .79, for example, suggests that the difference between the association and the importance weights lies mainly in the extraction method. The positive sign of the correlation indicates that the items which measure better higher-level concept are also rated as more important. Extended further, the finding can be translated into a hypothesis which states that the users' perceptions about paths to justice reflect both the actual experiences and the importance of the indicators used.

Although extracted from the same data set, the correlation between the FA and MR weights is lower than the association between the FA and SP weights[34]. Least correlated are the MR and the SP weights[35]. This is a surprising finding, given the high association between the FA and SP. Looking at the MR weights, it is apparent that the method produces somewhat difficult to interpret weights, which cannot be used without significant transformations (see Table 2: Weights). The discrepancy is a signal to critically scrutinize the potential of the multiple regression analysis, thus providing meaningful weights in the context of measuring access to justice.

Table 2: Weights

Indicator	Association weights - FA	Association weights -MR	Stated preferences weights
Decision control	0.81	0.37	0.86
Correctability	0.84	0.14	1.09
Bias Suppression	0.89	0.44	1.22
Process Control	0.76	0.04	0.99
Ethicality	0.77	0.05	0.85
Consistency	0.82	0.04	1.04
Accuracy	0.84	0.05	1.20

Figure 2 reveals the results from the application of the three different weights to the actual data. Visibly, the MR weighed data significantly deviate from the raw scores and the other two weighting options. The transformation is not only in the item scales,

---

[33] r=.79
[34] r=.60
[35] r=.13

but also in the relationships between the items. More consistency is significantly observed between the observed data and data weighted with the FA and the SP methods. When analyzing the relative position of the seven items, we find a negative correlation[36] between the observed data and the MR weighted data. Not only does the multiple regression method produce difficult to interpret results, but this method also significantly affects the ranks of the observed items.

Figure 2: Weighted data
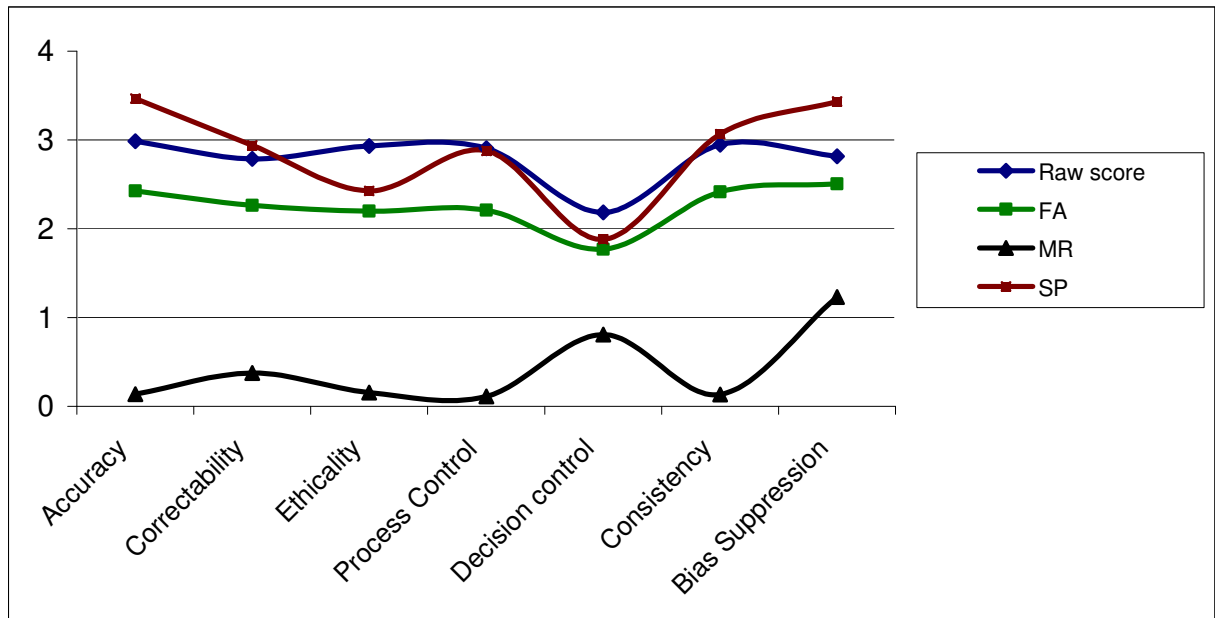
On the other hand, the FA and the SP methods change the order of the items to a much lesser degree in the observed data. The SP weighted data has a somewhat higher level of correlation with the observed data[37] than the FA method[38] does. Both correlations have similar strengths, and it could be concluded that the two methods produce comparable results. Another argument for the hypothesis that the association and the importance weights measure the same phenomena but with different approaches is the high correlation between the ranks of the items weighted with the two schemes[39].

The result of comparison between three models for extraction of weights suggests that the FA procedure has the most potential for generation of informative weights. Both theoretical and practical arguments imply the suitability of FA. First, the weights extracted from FA have intuitive meaning without any transformation. They represent the relationship of the item to the higher-order concept. Therefore, the weighting denotes that the items which are better related to the unobservable indicator will receive higher weights. From a practical point of view, the FA weights also have certain advantages. The weights could be extracted from the existing data without the need to ask the same sample of respondents new questions or to compute the weights from a different sample.

---

[36] r=-.58
[37] r=.57
[38] r=.43
[39] r=.96

Another major advantage of the FA weighting model is that it could be applied at higher levels in the Access to Justice Index. For instance, after extraction of the weights of the individual items, FA could be used to discover the weights of the sub-indicators into the indicators. In the next step, FA could be used to explore the weights of the three indicators before they are aggregated into a composite index.

Many of the advantages of the FA weights are inherent to the SP weights. Their most important benefit is the clear logic behind the stated preferences on the importance of the weights. An apparent disadvantage is the practical difficulties related to the collection of additional data that the method requires. As the analysis in the preceding paragraphs shows, after weighting with the FA and the SP weights, the data do not change dramatically. This could be an argument that the FA weights are a more efficient mean of weighting the observed data.

Weighting observed data could improve the validity, but only if the weights indeed adjust the data in the right direction. However, when certain requirements regarding the weights are not met, a non-weighted index will be superior to a weighted index. First, proper data must be obtained via a sufficiently large sample size. Second, regarding FA, the data could reveal negative factor loadings, which would cause a problem when applying and interpreting the weights to the data. Issues may also arise when the data do not meet expectations. An example would be the discovery of a multi-factor construct when the theory or the expectations suggest only one factor. In either of these two situations, the data should remain non-weighted. When one or more of these concepts is lacking, it has been suggested to refrain from using weights, as Nunnally has asserted that this option is often the best choice[40].

VI. Index validation

Once a composite index has been formulated to measure 'access to justice,' it is beneficial to validate the index, or, more specifically, focusing on external validation. External validation tests the validity of the index by examining the correlation of the final value and some other indicator of the variable under study. The current research uses the indicators discussed above to calculate the quality of the procedure (procedural justice, informational justice, interpersonal justice, and restorative justice) and the quality of the outcome (distributive justice, corrective justice, restorative justice, retributive justice, informational justice, utilitarianism, transformative justice, legal pragmatism, and formal justice). One can hypothesize that these indicators are measuring quality, or perceptions of satisfaction, with the procedure. As a result, the item that measures user satisfaction can serve as a presumed indicator measuring the same variable of quality of the procedure and quality of the outcome.

This index validation presupposes that those users who score a given path with a high rating are likely to state that they are satisfied with the procedure. Therefore, when scores are equivalent for this item and the index score overall, conclusions can be made regarding the accuracy of the survey. Specifically, it measures what it is supposed to measure. To briefly illustrate index validation of the Access to Justice Index, we will take the sub-index score of the quality of the procedure in the data

---

[40] NUNNALLY *supra note 32.*

from the consumer dispute commission study in the Netherlands, and examine its relation to the item measuring satisfaction with the procedure.

Table 3: Correlations Of Satisfaction With Non-Weighted Scores And Weighted Scores

Sub-indicator	Weighted Score using FA	Non-Weighted Score
Procedural Justice	.846**	.844**
Interpersonal Justice	.669**	.635**
Informational Justice	.645**	.646**

**. Correlation is significant at the .01 level (2-tailed).

The data conclude that while the difference is marginal, the weighted scores more accurately depict a valid index, assuming that satisfaction serves as a reasonable proxy variable. The question that does arise is whether this data is more in support of the non-weighting option, as the differences seem insignificant (i.e. .846 vs. .844). More importantly, however, is the strong correlation between the weighted scores and satisfaction, therefore suggesting the index is a valid measure of the quality of the procedure.

VII. Conclusion

The study and comparison of the cost and quality of paths to justice could be significantly facilitated by construction of composite index. On their ways to justice, the users evaluate distinct properties of the process. According to our theoretical framework, the three most important indicators are the costs, the quality of the procedure, and the quality of the outcome. These concepts were operationalized in detail in related articles and briefly outlined above.

Operationalization of the indicators, however, is only the first step in index construction. It is equally important to put in place a transparent and coherent scheme for rating the importance of the indicators. This scheme should be operationalized at the three levels at which indicators are aggregated. Establishing criteria for prioritizing the indicators is one of the most important phases of the index construction. Giving inappropriate weights to the indicators could tilt the composite index into a direction that compromises the validity and integrity of the measurement effort.

The importance of the indicators is expressed through the process of weighting. Not all observed indicators require re-assessment of their importance. Some indicators reflect factual statements, which could be true or false, but cannot be more important or less important. Other indicators, such as the quality of the procedure and the quality of the outcome, are subjective, meaning that two users of justice could assess similar experiences differently. Both assessments could be true, depending on the subjective value a person places on a particular indicator. It is completely possible that one user of justice values the opportunity to have her voice heard more than the possibility to correct wrong information. Another pair of users could disagree on what is more important in the outcome – its restorative effect or its enforceability. Based on the

subjective preferences, both users could be right. Apparently, the subjective preferences could have a strong influence on the composite index. If the preferences could be identified and the structure is stable, the observed data could be adjusted accordingly or, as otherwise said, weighted.

In the paper above, we discuss two models of weighting – stated preferences and extracted association weights. Despite the significant difference in the perspectives from which the two models extract the weights, we saw that when applied to empirical data both models produce very similar outcomes. Findings like this suggest that users of justice might have stable preferences with regard to the importance of the subjective indicators. However, we tested this hypothesis only with the quality of the procedure indicators. More research on other paths to justice is needed in order to formulate more conclusive inferences on whether there are universal, or at least path-specific, preference structures.

If both methods for weighting produce similar results, which one should be recommended? Before answering this question, we should consider the alternative of not using weights. Above, we discuss numerous disadvantages of the application of the weights. At least three groups of arguments against weighting and in favour of using the observed data could be outlined. First is the validity argument – weighting could adjust the data, but, at the same time, jeopardize its validity. In most studies, these risks will be difficult to control against, since there is often no base-line data to assess the performance of the index. Next, the practical argument that extraction of weights requires additional resources appears. This is particularly true for the importance weights. In the end, the application of weights should be critically reviewed against their impact on the data validity. In the text above, we outline a strategy for testing the validity of the weights against concepts, which could serve as proxies of the indicators (i.e. perceptions on satisfaction or favourability).

What was said in the preceding paragraph does not necessarily mean that weighting is harmful for the validity of composite index. It is a caution that only weights, based on solid data and sound theoretical models, should be used to adjust the observed data. If this is the case, we recommend the use of factor analysis for extraction of association weights. The model is based on the expectation that the lower-level items measure the unobserved upper-level items. In general, the association weights are easier to extract than the importance weights. In addition, our empirical data strongly suggests that the association weights extracted with factor analysis produce similar results to the importance weights.

With this analysis, the methodology for the Access to Justice Index has been reviewed and substantiated. The framework of the measurement has been outlined and the use of indexes was examined. The construction now allows for future evaluation on different legal paths a user encounters. As a result, not only can comparisons be made to provide the user with his or her best option, but improvements can be made where values depict an unsatisfactory picture. An accurately computed index is more likely to result by following the structure above and, subsequently, a greater understanding of the quality of justice will evolve.

Appendix A: Methodology for data collection

1. Data used for the extraction of association weights

Empirical data on experiences of users of justice in the Netherlands were collected through a cross-sectional survey in 2008. The study collects responses from individuals who experienced a vehicle-related consumer problem and sought resolution through the Dutch Consumer Dispute Commission. All individuals who addressed the Commission with vehicle-related problems in the previous 12 months received an invitation for participation in this study via the Consumer Dispute Commission. In this invitation letter, they were told that the purpose of the study was to assess their experience with the Consumer Dispute Commission concerning the cost and quality of the procedure and the outcome they obtained. The people who were invited were offered a monetary reward of € 25 for their participation.

In total, 850 individuals were invited to participate in the study. 230 of them responded to the invitation, resulting in a response rate of 27.1%. 34 of them were invited to participate in a focus group study. In addition, 74 respondents were asked to participate in the study at a later point in time, as it was decided to conduct a second survey with an improved version of the questionnaire. In total, 151 usable responses were collected.

2. Data used for the extraction of importance weights

In order to understand the users' preferences on the items of the quality of the outcome and quality of the procedure indicators, a quasi-experiment was carried out. 130 second-year law students from Tilburg University were invited to complete an online questionnaire that measured how the individual items are ranked against similar items[41]. For instance, the respondents were asked to rank the accuracy of the procedure against all 6 other indicators that measure procedural justice. In this group, the respondents had to rank the items from 1 to 7, with 1 representing the most important item and 8 representing the least important.

According to our theoretical framework, two or more indicators could carry equal levels of importance. Therefore, the questionnaire allowed placing the same ranks on more than one item. Hypothetically, a respondent could have chosen to rank all indicators as 1, which would have meant that they share equal levels of importance.

Four different types of the survey were distributed to groups of similar sizes. Each group was presented with a dispute case and its resolution, to which the students had to relate their responses[42]. The four cases had different underlying problems:
- Armed robbery;

---

[41] All law students enrolled in a course English Legal Language Skills were invited to participate. The questionnaire was distributed in English which might account for certain degree of measurement error.

[42] For the texts of the four case studies see http://www.measuringaccesstojustice.com/index.php/ELLS

- Relationship breakdown and split-up;
- Denial of student loan;
- Consumer problem with service provider.

In total, 102 valid responses were received, equivalent to a response rate of 78%. Particular reliability risks for the data collected through internet questionnaires were the two response order effects – primacy and yea-saying . In order to identify patterns in the data which signal those effects, we counted how many times a respondent selected identical values for two adjacent indicators. There are 33 indicators in the sample, grouped into 7 categories[43]. The mean value of the repeated scores is 8.11, with standard deviation of 5.95. A cut off value of 1.96 standard deviations above the mean[44] was selected for removing responses as plagued by primacy and/or yea-saying. Following this procedure, 4 responses were eliminated from the data set,[45] resulting in 98 total responses.

---

[43] 3 categories of indicators of the quality of the procedure, 3 categories of indicators of the quality of the outcome and 1 group which asks for the importance of the 3 general indicators – costs of the procedure, quality of the procedure and quality of the outcome.

[44] Cut off value = $19.77 \approx 20$ repeated rankings out of the 33 responses

[45] Additionally, the data set was analyzed for inconsistency in the ratings. The major concern was that for some set of indicators, non-consequential ranks were observed – i.e. item rated at 7th place but no item rated 6th. As this inconsistency would contradict the logic of the ranking, we re-organized the interrupted ranks into non-interrupted series. This transformation slightly decreased the mean ratings and improved the consistency of the data.

Appendix B: Computation of weights

1.   Association weights – principal components analysis

Table 4 reports the component loadings of the 7 items measuring the procedural justice concept. Only one component exceeds the eigenvalues threshold, indicating that the common variation is associated with one component. The component loadings are subsequently used as weights.

Table 4: Association weights – principal component analysis

Indicator	Component Loading
Accuracy	0.84
Correctability	0.84
Ethicality	0.77
Process Control	0.76
Decision control	0.81
Consistency	0.82
Bias Suppression	0.89

2.   Association weights - multiple regression analysis

Table 4 shows the un-standardized β coefficients from the multiple regression of the 7 items that measure the procedural justice on the satisfaction with the procedure. The β coefficients were used as weights without additional conversion.

Table 5: Association weights - multiple regression analysis

Indicator	MR
Accuracy	0.05
Correctability	0.14
Ethicality	0.05
Process Control	0.04
Decision control	0.37**
Consistency	0.04
Bias Suppression	0.44**

**The coefficient is statistically significant at the 0.01 level

3.   Importance weights

Unlike the association weights, the importance weights were obtained from a quasi-experiment with a sample of law students. The respondents were asked to report how important the particular indicator is with regard to a hypothetical dispute. Each indicator is ranked from 1 (most important) to n (least important, where n = the number of indicators). Therefore, lower ranks signified more important indicators. In

order to transform the ranks into weights, we first reversed the order of the indicator. From each score we extracted its deviation from the mean. The sign of the mean deviation was then reversed and again added to the mean. We take the natural logarithm of the reversed score as weight (see Table 6).

Table 6: Importance weights

	Original value	Reversed value	Natural logarithm
Accuracy	2.37	3.32	1.20
Correctability	2.71	2.98	1.09
Ethicity	3.35	2.34	0.85
Process control	3.00	2.69	0.99
Decision control	3.32	2.37	0.86
Consistency	2.88	2.82	1.04
Bias suppression	2.30	3.39	1.22