

ADVANCING THE STUDY OF TAX COMPLEXITY WITH THE USABILITY MODEL

*Frank H. Pedersen**

I.	INTRODUCTION	284
II.	CURRENT ANALYTICAL CONCEPTS IN THE FIELD OF TAX COMPLEXITY	293
	A. <i>The Concepts of Compliance Cost and Psychological Cost</i>	293
	1. Compliance cost according to Slemrod	295
	2. Psychological cost.....	297
	3. International Standard Cost Model.....	298
	B. <i>Notions of Tax Complexity</i>	301
	1. Diverse Notions of Tax Complexity	302
	2. Compliance Complexity, Transactional Complexity, and Rule Complexity	304
III.	CONFIGURATION OF THE CONCEPTS FOR ANALYSIS IN THE FIELD OF TAX COMPLEXITY	306
	A. <i>Context of Use</i>	308
	1. Users.....	310
	2. Tasks	311
	3. Interplay with Environment	312
	B. <i>Outcome of Use is Not Intrinsic to a Legal Authority</i>	314
	1. Description of Legal Authorities' Attributes	316
	2. Proxies and Causalities	318
	C. <i>Configuration of Elements in Assessments, and Underlying Meanings of "Tax Complexity"</i>	319
	1. Objectives of an Assessment	320
	2. Appraisal of the Assessed Result	321
	3. Possible Elements in an Assessment of Outcome of Use.....	323
	4. Tax Complexity Notions with Regard to Assessment of Outcome of Use	325
IV.	THE USABILITY CONCEPT.....	328
	A. <i>Definition of Usability</i>	330
	B. <i>Usability Measures</i>	334

* Frank H. Pedersen, Ph.D., Visiting Scholar, New York University; former Head, Tax Simplification Unit, Danish Ministry of Taxation. I am grateful to Deborah Schenk and Joshua Blank for their thoughtful suggestions and criticism throughout the drafting of this Article. I also thank Lily Batchelder, Graeme Cooper, Timothy Deering, Judith Freedman, Jeff Goodwin, Wolf Heydebrand, Kasper Hornbæk, Christopher Kotarba, Yoram Margalioth, Alex Raskolnikov, Søren Riis, Wolfgang Schön, Daniel Shaviro, Richard Stewart, and Tetsuya Watanable, as well as participants in the NYU Law Global Fellows Forum and NYU Economic and Political Sociology Workshop, for helpful discussions and comments on early drafts. Kirstin Peterson provided excellent editorial assistance. Financial support from among other the Hauser Global Law School Program and the Study and Research Foundation of The Institute of State Authorized Public Accountants (FSR) in Denmark is gratefully acknowledged. All mistakes are solely my own.

1.	The Effectiveness Measure	335
2.	The Resource-Efficiency Measure	337
3.	The Satisfaction Measure.....	338
C.	<i>Objective and Subjective Measures</i>	341
D.	<i>Usability Improvements</i>	345
1.	Appraisal of Assessed Level of Usability.....	345
2.	Focus of Usability Improvements	347
E.	<i>Low or High Usability cf. Complex</i>	350
V.	ENHANCEMENT OF COMPREHENSION OF WHAT HAPPENS FOR TAXPAYERS	351
A.	<i>More Suitable Tools Than Current Notions of Tax Complexity</i>	351
B.	<i>Confinements by Objectives: Computation of Tax Liability as Illustration for Summing Up</i>	356
VI.	CONCLUSION.....	360

Abstract

Taxpayers, business employees, and tax professionals perform various tax compliance tasks, such as finding relevant information, filling out forms, and performing computations – perhaps with the help of tax-preparation software. Carrying out such activities can result in defrayed cost, time consumption, uncertainty, frustration, and the like. The resulting resource-spending, level of psychological well-being, and degree of achievement in completing the attempted compliance tasks have welfare implications. Improvements in taxpayers' welfare relating to compliance tasks will come from amelioration of negative occurrences such as taxpayer frustration. It follows that the more thoroughly taxpayers' encounters with compliance tasks are analyzed, the more effectively policymakers can work toward mitigation of unwanted occurrences. This Article introduces the usability model as a tool for such analysis.

Usability is a well-established theoretical model for analyzing human performance of tasks, but it is a novel concept in the legal context. At the heart of usability is recognition of people's desire to experience positive feelings of success, competence, and mastery in the performance of tasks. As a tax policy objective, usability emphasizes taxpayers' aspiration to work smoothly as they perform the activities necessary to determine their tax liability. Because usability goes beyond economic cost to measure degrees of task accomplishment and

psychological well-being in response to task performance, analyses based on usability are wider-ranging than those based on the compliance cost concept.

The current literature is generally limited in using the term "tax complexity" to describe both tax rules and the way compliance tasks are carried out. This Article demonstrates that "complexity" is too narrow because it focuses on tax rules in isolation from the ways taxpayers actually carry out their compliance tasks. In addition, current notions of tax complexity are insufficient to provide guidance to policymakers because those notions are too vague in establishing which kind of taxpayer occurrences they embrace. Still, examination of the diverging current notions of tax complexity facilitates a more complete understanding of the field of tax complexity. Therefore, this Article discusses how various existing complexity concepts for assessment, such as compliance cost, relate to each other by identifying the issues each of them addresses, as well as how those issues are encompassed by the usability model. This helps to clarify why policymakers planning improvement initiatives would gain more traction by reasoning along the lines of usability, rather than attempting to apply vaguely-defined notions of complexity.

The beginning of wisdom is to call things by their right names.

-Chinese Proverb

I. INTRODUCTION

This Article attempts to advance the field of tax complexity by shedding new light on the main concepts and notions upon which the field relies, and, most significantly, by presenting a new, more comprehensive analytical concept for studying the experience of taxpayers (and others who do work on their behalf) in performing tax compliance tasks.

To comply with income tax law, taxpayers perform activities, such as information retrieval and computation, in furtherance of broader tax compliance tasks such as preparing tax returns for filing. Deriving from the performance of activities as part of broader compliance tasks are various empirical occurrences: taxpayers spend time and defray costs, they might become agitated or exhausted, and they might not complete their tasks to the degree they intended. The occurrences encountered by taxpayers as they attempt to complete tax-related tasks have welfare implications, which lead to policy aspirations for improvement in how taxpayers perform. Consequently, detailed

assessment and comprehension of the empirical occurrences commonly deriving from taxpayers' performance of compliance activities – occurrences such as time consumption or unintentional divergence from legally correct compliance – are important to enable policymakers to ascertain whether they see the specific occurrences as representing problems which should be mitigated.

For taxpayers, this Article argues, progress in compliance efforts is all about amelioration of undesirable empirical occurrences brought about by the carrying out of activities for particular compliance tasks. Hence, conceptual advances in assessing what taxpayers encounter in performing tax-related tasks will allow policymakers to better fulfill the aspiration for improvements by tailoring initiatives more precisely toward encounters which are appraised as unacceptable. For example, where the assessed level of frustration deriving from a particular compliance task is found to be too high, reducing frustration will be the goal of mitigation efforts.

Consideration of taxpayer frustration, time consumption, and other occurrences, which are caused by compliance efforts, takes place in the context of discussions of tax complexity. Tax complexity is one of the three traditional grounds upon which tax policy is evaluated; the other two are efficiency and equity.¹ However, even though tax complexity is a crux of tax policy evaluation, there is concurrence within the literature that precisely what is meant by “tax complexity” and cognate terms is unclear,² and that by extension, the terms “complexity” and “simplification” do not contribute appreciable clarity to the field

1. See, e.g., MICHAEL J. GRAETZ & DEBORAH H. SCHENK, *FEDERAL INCOME TAXATION: PRINCIPLES AND POLICIES* 28 (6th ed. 2009) (“There is widespread agreement that the criteria to be used in evaluating taxes are equity, efficiency, and simplicity. There is considerably less agreement, however, as to the precise meaning of these criteria . . .”).

2. See, e.g., Graeme S. Cooper, *Themes and Issues in Tax Simplification*, in 2 *TAXATION: CRITICAL PERSPECTIVES ON THE WORLD ECONOMY* 238, 240 (Simon R. James ed., 2002) (“In short, simplification is not one debate but many masquerading under a common label.”); Sidney I. Roberts, *The Viewpoint of the Tax Adviser: An Overview of Simplification*, *THE TAX ADVISER* 32, 34 (Jan. 1979) (recognizing the vagueness of the words “simplification” and “complexity” in the tax literature, and remarking: “It would be salutary if the Conference forbade the use of the terms complexity and simplicity. Those terms have too many meanings; and barring their use would insure the discussion relates to the same subject, at least at the same time.”); Stanley A. Koppelman, *At-Risk and Passive Activity Limitations: Can Complexity be Reduced?*, 45 *TAX L. REV.* 97, 98-101 (1989) (arguing about the consequences of discussions of complexity that rarely attempt to define what complexity means); Edward J. McCaffery, *The Holy Grail of Tax Simplification*, 1990 *WIS. L. REV.* 1267, 1269 (1990) (“It is not easy to arrive at ready definitions of ‘simplicity’ and its cognates and antonyms.”).

but are better understood for now as labels for the field.³ Yet, in contrast to earlier work in the field, this Article does not see the essential challenge as reaching an overall definition of “tax complexity” *per se*. Instead, this Article takes the position that what is crucial for increasing the value of the field’s scholarly studies is for the literature to be exact about which underlying phenomena references to the term “tax complexity” actually embrace, for instance, whether “tax complexity” designates certain special tasks or particular empirical occurrences, or whether it refers to characteristics embedded in statutes.

This Article demonstrates that questions about how well taxpayers can perform necessary activities in completing their tax compliance tasks⁴ – whether compulsory tasks such as reporting income, or optional tasks such as obtaining deductions – cannot be answered by examining tax statutes and other legal authorities⁵ independent of the real-world context in which taxpayers actually execute their tasks. Analyses, which are limited to descriptions of legal authorities, examine phenomena that are distinct from the empirical occurrences that derive from the performance of tax-compliance activities. Hence, any increased complexity that may exist in the legal basis for the income tax⁶ does not necessarily justify the conclusion that taxpayers’ performance of tax-related activities will deteriorate

3. In fact, the contours of the field of tax complexity and simplification appear to arise from whatever is not included by the two other criteria for tax policy, *i.e.*, efficiency and equity. As a field, tax complexity thus does not engage issues of tax liability and tax fairness, which are encompassed by the two other grounds of evaluation. *See generally infra* Part II.A.(quoting definitions of compliance cost). Hence, inquiries within the field as to what taxpayers encounter are concerned only with what occurs as they carry out activities, not with the resulting tax liability.

4. Throughout the Article, “task” generally refers to what a taxpayer is attempting to achieve, for example, fulfilling an informational reporting obligation; “activities” refers to all the actions, which are undertaken in order to execute the task, for example, collection of data or computation of an amount. “Occurrences,” “aspects of performance,” and “what taxpayers encounter” all refer to empirical consequences, such as time consumption or monetary expenditure, which derive from performance of the activities to fulfill the task. *See generally* McCaffery, *supra* note 2, at 1271-1317.

5. This Article uses the term “legal authority” as Deborah Paul has described it: “I use the term ‘authority’ or ‘legal authority’ to refer to a rule, standard, or other norm with legal weight. For example, the Internal Revenue Code and Treasury Regulations contain many authorities. A judicial opinion may contain one or more authorities.” Deborah L. Paul, *The Sources of Tax Complexity: How Much Simplicity Can Fundamental Tax Reform Achieve?*, 76 N.C. L. REV. 151, 157 n.17 (1997). *But see* Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557, 559 n.2 (1992) (using the phrase “legal commands”).

6. The question whether the income tax, understood as a confluence of legal authorities, is becoming more and more complex is one that this Article understands as being answerable in quantifiable terms, at least in principle, based on a suitable definition of complexity. *See generally infra* Part III.B.1.

as a result.⁷ It is even possible for taxpayers' performance to improve, for example, with a decrease in time consumption, despite increasing complexity in the legal authorities. This fact explains the schism in recent decades between complaints about increased statutory complexity, on the one hand, and, on the other, taxpayers' experience of the burden of tax-related tasks being lightened somewhat by their use of tax-preparation software.⁸ This divergence is a development, which the tax literature hitherto has not explained in a theoretically coherent way.

To provide adequate guidance for policymakers' initiatives to ameliorate undesirable occurrences deriving from taxpayers' performance, a notion of "tax complexity" would have to pertain to taxpayers' empirical experiences in handling their tax affairs and would have to elucidate such occurrences in enough detail to establish the specific problem to be mitigated. For instance, an analysis would not be sufficient if it merely asserted that some tasks, such as recordkeeping requirements, are problematic, without having assessed the empirical occurrences, like time consumption and level of frustration, that derive from executing those tasks in a particular context. This Article demonstrates the limitations of existing notions of tax complexity in satisfying this requirement.

Moreover, the Article argues that much of the guidance which current tax complexity notions are incapable of providing in this regard can be gained instead through analyses employing the "usability" model, which this Article introduces as an innovation in the field of tax complexity. Usability is a well-developed concept concerned with "quality in use."⁹ A high level

7. Bayless Manning offers a colorful statement about such a decline: "Measured by any and every index, our law is exploding. . . . Statutory codes, such as those in the fields of commercial law and taxation, are becoming ever more particularistic, longer, more complex, and less comprehensible. We are drowning in law." Bayless Manning, *Hyperlexis: Our National Disease*, 71 NW. U. L. REV. 767, 767 (1977). Interestingly, however, the concern Manning expresses in his shrewd and visionary article appears to pertain mainly to how well taxpayers are able to perform.

8. See, e.g., Randy Johnston, *Easier Tax Preparation With Automation*, CPA PRACTICE ADVISOR July 1, 2009, <http://www.cpapracticeadvisor.com/article/10268307/easier-tax-preparation-with-automation> ("Most firms reported a very smooth and profitable tax season this year. I hope your firm had this type of experience. We believe that current tax preparation software is sophisticated, does most of the job correctly, and accommodates many oddities of state and local tax. Not everything is perfect, but the products are very stable and comprehensive from the high end down to the low end.").

9. See, e.g., Kasper Hornbæk, *Current Practice in Measuring Usability: Challenges to Usability Studies and Research*, 64 INT'L J. HUMAN-COMPUTER STUDIES 79, 86 (2006) (applying the designation "quality-in-use" as a simplified description of what usability

of usability is present in the performance of a task when the person who executes the task is successful in carrying it out smoothly in its real-world manner and context, without the performance causing discomfort.¹⁰ Although usability is a novelty in the legal context, it is an established theoretical model for analyzing people's performance of tasks in many other contexts.¹¹ This Article employs the specific usability model promulgated by the International Organization for Standardization (ISO), whose paradigm is the most widely acknowledged model of usability.¹²

A usability assessment in the income tax context could concern a particular segment of people (or "users")¹³ performing

embraces); *see also* Quality management, *infra* note 12 (referring to the corresponding phrase "quality management").

10. *See infra* Part IV and IV.A.

11. *See* Abran, *infra* note 20.

12. The extraordinarily influential standard of the International Organization for Standardization ("ISO"), namely ISO 9241, serves as the starting point for many other usability concepts. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, ERGONOMIC REQUIREMENTS FOR OFFICE WORK WITH VISUAL DISPLAY TERMINALS (VDTs) – PART 11: GUIDANCE ON USABILITY (1998) (hereinafter "ISO 9241-11"). *See, e.g.*, BEN SHNEIDERMAN & CATHERINE PLAISANT, DESIGNING THE USER INTERFACE 14 (5th ed. 2010) (noting ISO's breadth); Hornbæk, *supra* note 9, at 80 (taking ISO's usability concept as the grounding for his alternative). Its prominence makes it the obvious choice for an exploratory article such as this.

ISO is a non-governmental organization composed of a network of national standards institutes from 162 countries, with one member per country, and a Central Secretariat in Switzerland. International Organization for Standardization, *About ISO*, <http://www.iso.org/iso/about.htm> (last visited Oct.10, 2011). It is the world's largest developer and publisher of international standards. *Id.* International standards in general provide specifications and criteria to be applied consistently in the provision of services, the manufacture of products, and the classification of materials. *See* International Organization for Standardization, *General Information on ISO*, http://www.iso.org/iso/home/faqs/faqs_general_information_on_iso.htm (last visited Oct.10, 2011).

This Article draws upon ISO 9241-11, which introduces the concept of usability but does not make specific recommendations of product attributes; rather, it pertains solely to the benefits of measuring usability. In this way, ISO 9241-11 is quite different from most standards, which do include advice on product attributes. *See* International Organization for Standardization, ISO 9000 - Quality management, http://www.iso.org/iso/home/standards/management-standards/iso_9000.htm (last visited Oct.10, 2012) ("The ISO 9000 family addresses various aspects of 'Quality management.'") [hereinafter "Quality management"].

13. This Article refers to the persons fulfilling a task induced by a legal authority as "taxpayers" or "users," and applies those terms broadly, so that they also embrace employees and professionals working on behalf of those legally liable for paying taxes. For expositional ease, this Article generally uses the terms "taxpayer" and "user" without any appreciable distinction. If the person – or legal person – who bears tax liability does not perform an activity, but, for instance, purchases assistance from a tax preparer or employees, then the person who bears tax liability is not the user with regard to activities that she or he does not perform. Level of usability can be assessed just as well for activities, which tax preparers or employees of businesses carry out, and in those cases,

the activities required to complete the task of obtaining a deduction for charitable contributions.¹⁴ Such an assessment would measure empirical occurrences showing: the degree to which this segment can accomplish the activities required to complete the task; their resource-spending; and the degree to which they can perform the task without anxiety. Together, these measures comprise the level of usability of the performance.¹⁵ If the usability level were appraised as too low, the analysis would show in detail which occurrences should be the focus of improvement initiatives.

In introducing the usability model, this Article distinguishes between the usability *framework* and the usability *concept*. Whereas the usability *framework* offers general principles to follow to ensure validity in assessments of task performances,¹⁶ the usability *concept* provides a particular approach to measuring and comprehending people's performance of tasks in terms of level of usability.¹⁷ This Article argues that principles from the usability *framework* are important not only for usability assessments, but also for ensuring validity in assessments based on other concepts that may be applied to assess empirical occurrences deriving from taxpayers' execution of compliance-related activities. The most prevalent of these concepts is compliance cost, which measures aspects of resource-spending in terms of economic cost.¹⁸ Another existing concept for evaluating some aspects of taxpayers' experience is "psychological cost," which assesses occurrences such as stress, anxiety, and the like resulting from tax compliance activities.¹⁹

Legal scholars have not previously applied either the usability concept²⁰ or the usability framework in the legal

they are considered to be the users. The Article does not contemplate usability for tax officials or others who do not act in the capacity of taxpayers or on the taxpayers' behalf.

14. The example of a deduction for charitable contributions will recur throughout the Article, as it is easily understandable and is common to many countries' income tax systems. See HUGH J. AULT ET AL., *COMPARATIVE INCOME TAXATION: A STRUCTURAL ANALYSIS* 239-42 (3rd ed. 2010) (reviewing the charitable deduction for eight countries).

15. See sources cited *infra* note 192.

16. See *infra* Part III.

17. See *infra* Part IV.

18. See *infra* Part II.A.

19. See *infra* Part II.A.2.

20. The usability concept is commonly applied to evaluate user performance which employs information and communication technologies – for example, people's performance when employing LexisNexis to find an article.

The usability concept, however, may be extended beyond its normal application. See, e.g., Alain Abran et al., *Usability Meanings and Interpretations in ISO Standards*, 11 *SOFTWARE QUALITY J.* 325, 331 (2003) (recognizing the breadth of the ISO usability

context.²¹ This Article contends that usability is an essential and valuable complement to existing methods of assessment, and that reasoning along usability lines can benefit policymakers in planning improvement initiatives.²² The Article argues that usability can more comprehensively assess what occurs in people's performance of tax-related tasks than any of the concepts currently being applied. Yet usability is offered as a supplementary concept, as each concept's manner of assessment – whether as compliance cost or as level of usability – complements the others in elucidating taxpayers' empirical experience, and relates to different tax policy objectives.

When increased usability is the objective, the occurrences highlighted as possibly needing amelioration are different from those focused on when the policy goal is merely to minimize compliance cost. Usability emphasizes people's attainment of a high level of achievement in executing tasks while being at ease in their performance.²³ Usability also emphasizes the relevance of taxpayers' own perceptions of their performance, rather than a purely objective assessment of what has occurred.²⁴ For instance, while a user segment might attain a high level of achievement as a factual matter, these users might nonetheless believe, subjectively, that they have failed to achieve even a modest level. Assessment of taxpayers' perceived occurrences can detect such a discrepancy and make it possible to aim improvement initiatives at ameliorating taxpayers' subjectively negative experiences when policy objectives include this.

standard: "In spite of the name, the definitions in part 11 [of ISO 9241] are also known to be applicable to other situations where a user interacts with a product to achieve certain objectives. This extension makes usability a generic usability concept, likely applicable outside its conventional applications in information technology").

21. To the Author's knowledge after a thorough literature review, the application of the usability model to legal work in the manner this Article proposes is entirely novel. Although analyzing processes is a recognized application area for ISO 9241-11, some adjustment of the concept has been necessary for the present application. As such, references to the ISO standard are not intended to be an invocation of its authority. In addition, ISO 9241 has a reputation for being inaccessible and its exact understanding has been contested. *See, e.g.*, David Travis, *Bluffers' Guide to ISO 9241 4 (Bluffers' Guide to Usability Standards 2009)*; Hornbæk, *supra* note 9, at 96-97. A discussion of the different understandings and concepts of usability is beyond the scope of this Article.

22. It is important to note that the usability concept is likely also suitable for assessing performance of activities for legal tasks in other areas of the law, in addition to income tax. However, for expositional ease, this Article focuses only on income tax issues. In addition, since this is a conceptual article, the discussion is not specifically geared to one country's income tax system, or to the U.S. federal income tax as opposed to state or local income tax.

23. *See infra* Parts IV.B.1., and IV.B.3.

24. *See infra* Part IV.C.

In the field of tax complexity, “complex” is currently applied both to describe legal authorities (tax rules) and to describe assessed empirical occurrences that derive from taxpayers’ performance of tax-related tasks.²⁵ Usage of the same term to denote such distinct phenomena, aggravated by the terms having been applied to designate other, more specific matters as well, results in a great risk of ambiguity and misunderstanding. To help rectify this, this Article presents, as an additional contribution, an analysis pinpointing the diverse underlying issues with which each current notion of tax complexity is concerned. This analysis supports the argument that the term “tax complexity” should be explicated with reference to the concrete phenomena, it is meant to include. If “tax complexity” is intended to pertain to what taxpayers encounter in the course of their performance, then the empirical occurrences to be embraced should be clarified. The Article claims that the term “tax complexity” should not continue to be accepted as a seemingly indefinable hovering cloud.

The pervasive use of tax software is the most ubiquitous recent alteration of the context in which tax-related tasks are performed.²⁶ The usability framework can put into perspective how the compliance burden has been lightened by new information and communications technology.²⁷ The usability model can thus provide a theoretically coherent way to resolve empirical issues such as the apparent contradiction between indicia of increased tax-law complexity and the concomitant decrease in the burden of tax compliance, a phenomenon which has been noted by taxpayers and attributed to tax-preparation software.

25. See *infra* Part II.B.

26. See e.g. Lawrence Zelenak, *Complex Tax Legislation in the TurboTax Era*, 1 COLUMBIA J. TAX LAW 94, 94-95 (2010) (noting that in 1987 only 13% of all individual returns had been prepared on computers, growing to 67% for tax year 1997 and 89% for tax year 2006. These numbers include both paid preparer returns and self-prepared returns, of which the percentages prepared on computers for tax year 2006 were 98% and 71%, respectively). For the development of technology application internationally, see OECD, CENTRE FOR TAX POLICY AND ADMINISTRATION, FORUM ON TAX ADMINISTRATION, TAX ADMINISTRATION IN OECD AND SELECTED NON-OECD COUNTRIES: COMPARATIVE INFORMATION SERIES 2010 186 (March 2011) (hereinafter “COMPARATIVE INFORMATION SERIES”)(describing that there has been substantial progress in recent years in the number of revenue bodies offering electronic filing capabilities for the major taxes and that “Just over half of revenue bodies reported that the majority of their personal taxpayers used e-filing (either by themselves or via tax professionals) for the 2009 year.”).

27. See *infra* Part III.A. (explaining how the weight of a specific compliance burden deriving from a task depends on how it is being undertaken, for instance, with or without the help of tax software); see also *infra* Part V.B.

As a conceptual piece,²⁸ this Article's intention is not to provide examples of assessments²⁹ or specifications for how to conduct usability evaluations,³⁰ nor does it advocate specific initiatives. Rather, its goal is to contribute to theory development in the field of tax complexity and simplification, in particular, by introducing the usability model as an innovation with important ramifications for advancement of the field.

The remainder of this Article is presented as follows: Part II reviews extant concepts available in the field of tax complexity for analysis of what taxpayers encounter when performing tax-related tasks. Part III introduces the usability *framework* and demonstrates a way to configure those extant concepts for analysis, based on the usability framework's general principles for analysis of human task performance. Part IV presents the usability *concept*, which specifically assesses usability, or "quality-in-use." Part V discusses the advantages of applying the usability model, and identifies tax policy objectives to which the usability concept has relevance. A discussion of computation of tax liability sums up. Part VI concludes the Article.

28. "Conceptual article" is understood here as one that focuses primarily on theory development and does not present data or analysis for purposes of theory testing. In the spectrum of conceptual articles, this Article could be characterized as concerning what has been termed "discovery." See generally Manjit S. Yadav, *The Decline of Conceptual Articles and Implications for Knowledge Development*, 74 J. MARKETING 1, 2 (2010) ("In general, the context of discovery is related to the conception of new ideas (e.g., new constructs) or to the creative synthesis of existing ideas (e.g., new relationships between well-accepted constructs)."). Work done towards discovery is distinct from work done towards "justification"; "The context of justification is the realm in which data and analytical procedures are employed to establish the plausibility and acceptability of these ideas." *Id.*

29. The value of simply reasoning about usability's importance is addressed by Jacob Nielsen, who writes:

Only by defining the abstract concept of "usability" in terms of more precise and measurable components can we arrive at an engineering discipline where usability is not just argued but is systematically approached, improved, and evaluated (possibly measured). *Even if you do not intend to run formal measurement studies of the usability attributes of your system, it is an illuminating exercise to consider how its usability could be measurable.*

Jacob Nielsen, *USABILITY ENGINEERING* 26 (1993) (emphasis added).

30. An advantage of drawing on a concept that is already in use, even if not in a legal context, is that for purposes of further research in conducting usability evaluations, some knowledge is already available for reference in working out the particulars that must be concretized before evaluations can occur. For those conducting assessments of compliance cost, there is a comprehensive manual, which focuses on the International Standard Cost Model (hereinafter "SCM"). SCM NETWORK, *INTERNATIONAL STANDARD COST MODEL MANUAL: MEASURING AND REDUCING ADMINISTRATIVE BURDENS FOR BUSINESS* 8-9 (2005) (hereinafter "SCM MANUAL"). The development of such manuals is an extensive project in itself, but a manual for usability could be based on the general accumulated experience in conducting usability assessments. The SCM approach to assessment of compliance cost is discussed *infra* Part II.A.3.

II. CURRENT ANALYTICAL CONCEPTS IN THE FIELD OF TAX COMPLEXITY

This Part reviews prevalent concepts and notions employed in the tax complexity and simplification literature for analysis of what taxpayers encounter, as an empirical matter, in the performance of their tax compliance-related tasks. Assessments made under each of these concepts include divergent aspects of what taxpayers encounter. If, for instance, a taxpayer spends ten minutes carrying out the activity of computation, a compliance-cost study would assess the occurrence of ten minutes' time consumption. By contrast, an analysis using the concept of psychological cost would assess different empirical occurrences, such as stress and anxiety resulting from the computation activity. Section A of this Part considers the focus and methods of the compliance cost and psychological cost concepts. A careful consideration of the different empirical occurrences assessed by each concept, and the different ways in which the concepts measure and comprehend those occurrences, will demonstrate the critical differences between the usability concept and the other concepts regarding the ways in which they elucidate taxpayers' experiences and, by extension, aid policymakers in developing improvement initiatives.

In addition to compliance cost and psychological cost, the tax literature also employs notions it labels as "complexity" in the analysis and discussion of what taxpayers experience in executing tasks.³¹ However, it is often unclear exactly what constitutes this tax complexity.³² Thus, with regard to taxpayers (or others who perform tax-compliance activities on their behalf), the question is: What are the empirical occurrences being labeled "tax complexity"?³³ Section B of this Part describes various notions of tax complexity, which are generally presumed to be similar to compliance cost and psychological cost, in the sense that they also pertain to analysis of what taxpayers encounter empirically.

A. *The Concepts of Compliance Cost and Psychological Cost*

The compliance cost concept is by far the most prevalent tool for analyzing what occurs for taxpayers in the performance of

31. See *infra* Part II.B.

32. See *infra* Part II.B. and discussion in Parts III and V.

33. See discussion *infra* Part III.C.4. and Part V.A.

tax-compliance tasks.³⁴ One definition of compliance cost comes from Sandford, Godwin, and Hardwick: “cost incurred by taxpayers or third parties, notably business, in meeting the requirements laid on them by a given tax structure (excluding the payment of the tax itself and any distortion cost arising from it).”³⁵ When referred to generally in this Article, “compliance cost” is to be understood in accordance with this definition. In practice, compliance cost comprises not one but several concepts,³⁶ whose common denominator is that they assess occurrences deemed resource-spending which can all, in some way, be understood as “economic cost.”³⁷ Some differences among the varying concepts pertain to whether they focus on assessment of compliance cost at an aggregate level, or on obtaining detailed information that may be more useful in making granular improvements.³⁸ The two compliance-cost concepts discussed in Subsections 1 and 3 below illustrate this difference in focus. Psychological occurrences that derive from compliance tasks, such as stress and anxiety, are understood as

34. See, e.g., OECD, CENTRE FOR TAX POLICY AND ADMINISTRATION, FORUM ON TAX ADMINISTRATION: TAXPAYER SERVICES SUB-GROUP, INFORMATION NOTE, PROGRAMS TO REDUCE THE ADMINISTRATION BURDEN OF TAX REGULATION (follow-up report, March 2010) at 9 (hereinafter “FORUM ON TAX ADMINISTRATION”)(summarizing recent developments in reducing administrative burden and compliance cost).

It is worth noting how gradually the concept of compliance cost achieved the prevalence it enjoys today, and what steps it underwent to do so. Jeff Pope argues that the following six periods characterize its development:

First, there is the period in which the compliance cost of taxation is unrecognized and hidden; there is no interest in the topic at all, other than perhaps the occasional conceptual or theoretical acknowledgement . . . The second phase may be termed “professional, qualitative recognition” . . . The third phase is that of “estimation and evaluation” . . . The fourth phase in the development of compliance costs may be identified as “policy recognition” . . . The fifth phase is that of “effective policy measures,” in the sense that not only are policy, legislative and operational changes introduced, but they are actually effective in reducing the compliance cost of taxation at a grass-roots level. The sixth and final phase is the continual monitoring of compliance cost, as well as administrative costs. This should include published compliance and administrative cost assessment for every major tax change.

Jeff Pope, *The Compliance Costs of Taxation in Australia and Tax Simplification*, 18 AUSTRALIAN J. MGMT. 69, 71 (June 1993)[hereinafter *Tax Simplification*].

35. CEDRIC SANDFORD ET AL., ADMINISTRATIVE AND COMPLIANCE COST OF TAXATION xi (1989).

36. See, e.g., FORUM ON TAX ADMINISTRATION, *supra* note 34, at 9 (explaining some of the different designations which are applied: “[t]he term administrative burden is often used interchangeably with other terms including: ‘red tape’, ‘compliance burden’, or ‘compliance costs’.”).

37. See discussion in this Section.

38. The SCM approach emphasizes that “the results from the SCM measurements are directly applicable to governments’ simplification work, in that the results show the specific regulation and its details which are especially burdensome for businesses.” SCM MANUAL, *supra* note 30, at 8. See further discussion, Section 3.

being covered by compliance cost in principle.³⁹ In practice, however, such aspects of psychological well-being are largely omitted from assessments because of the difficulty of calculating them in economic terms.⁴⁰ Instead, the tax literature discusses such aspects separately as “psychological cost.”⁴¹ Subsection 2 describes the concept of psychological cost.

1. Compliance cost according to Slemrod

Joel Slemrod designates compliance cost as “complexity,” and argues that a useful measure of complexity is “the total resource cost of collecting the taxes, which is the sum of the tax collection agency’s budget, the value of the time and money spent by taxpayers, and any cost incurred by third parties in the collection process.”⁴² He regards changes in total resource cost as changes in tax “simplicity” or tax “complexity,” terms that become antonyms in his work.⁴³ Slemrod emphasizes that he is adopting a particular interpretation of what constitutes simplicity in a tax system, by characterizing simplicity according to the value of the resources expended in complying with the law and enforcing it.⁴⁴ Hence, this concept focuses on assessments of all the costs of operating the tax system; a cheaper operation means “simpler,” and a more costly operation means “more complex.”⁴⁵

“This measure,” Slemrod states, “had the advantage of being quantifiable, although not without error, but in some cases may conflict with intuitive ideas about complexity.”⁴⁶ In some of his early work, Slemrod presented examples illustrating how he perceived that his measure might seem to be at odds with an intuitive concept of simplicity:

39. See Subsection 2.

40. See Subsection 2.

41. See Subsection 2.

42. Joel Slemrod, *The Etiology of Tax Complexity: Evidence from U.S. State Income Tax Systems*, 33 Public Finance Review 281 (2005) [hereinafter *Etiology*]. Whereas Slemrod’s definition includes the tax revenue body’s budget, Sandford et al.’s definition of compliance cost does not include what they term “administrative cost: public sector costs incurred in administrating an existing code (including advice on its modification).” SANDFORD ET AL., *supra* note 35, at xi.

43. Joel Slemrod, *Optimal Tax Simplification: Toward a Framework for Analysis*, in NAT’L TAX ASS’N – TAX INST. OF AM. 158 (1984) (hereinafter *Optimal Tax Simplification*) (“I will characterize a tax system’s simplicity by the value of the resources that are expended in complying with the law and enforcing the law.”).

44. *Id.* at 159.

45. *Id.* at 158.

46. *Etiology*, *supra* note 42, at 281.

The General Accounting Office has estimated that in 1976 over six million taxpayers who should have filed income tax returns did not do so. In finding that the average educational level of the non-filers was below the national norm suggested to them that the reason for non-filing was often the process was too complicated to be understood. Assuming the IRS makes no effort to uncover the non-filers, the collection cost associated with this group is approximately zero. Furthermore, a change in the tax system which, by making the process easier to comprehend, enables those currently not filing to file would almost certainly add to total collection costs. Although the tax system may have become simpler to understand, it has become more complicated (or, specifically, more costly) to collect the revenue. Our cost-based measure of simplicity embraces the latter judgment.⁴⁷

Slemrod's example, besides illustrating the semantic challenges that arise in the terminology of complexity and simplification, explicates the boundary between the occurrences that a compliance-cost assessment includes, and those it excludes. Consistent with its objectives, a compliance-cost study does not examine whether taxpayers can actually accomplish their tasks.⁴⁸ In the example, this is the aspect of taxpayers' encounters, which has improved with the hypothetical change enabling more taxpayers to file. Another aspect, which the compliance-cost concept leaves unexamined, is users' response to the process of managing the task, where it is likely that users' level of comfort is higher when they are actually able to accomplish a task they attempt.⁴⁹

In short, the compliance-cost concept assesses occurrences of resource-spending, which are monetary, like defrayed cost, or which have traditionally been transformed into monetary terms, such as time consumption.⁵⁰ Although, stress and similar psychological occurrences are also understood in theory to be costs of compliance, the difficulty of computing such costs in

47. *Optimal Tax Simplification*, *supra* note 43, at 159.

48. *See, e.g.*, the definition by SANDFORD ET AL., *supra* note 35, Part II (encompassing "cost" but not task completion).

49. *See* discussion *infra* Part IV.B.3.

50. *See Etiology*, *supra* note 42 and accompanying text (defining compliance cost).

monetary terms results in the relegation of those costs to a category by itself, designated as psychological cost.⁵¹

2. Psychological cost

In principle, the concept of psychological cost is available for analysis of what taxpayers performing tax-related tasks encounter that is related to psychological well-being. However, it is rarely applied.⁵² Despite the fact that scholars have recognized undesirable psychological occurrences deriving from taxpayers' performance since the work of Adam Smith,⁵³ the tax literature addressing psychological occurrences is still extremely limited.⁵⁴ The most thorough discussion seems to be by Sandford, who explains what comprises psychological cost as follows:

[D]ifficult or impossible to measure satisfactorily[,
but] an important component of compliance cost.
Many people experience considerable anxiety and

51. In his original definition of total resource cost, Slemrod also included "the value of the uncertainty and displeasure [taxpayers] experience in filing their returns." *Optimal Tax Simplification*, *supra* note 43, at 159.

52. Joseph Bankman expresses the general understanding that appears to underlie this situation: "Economists have attempted to measure the time spent on filing and fees paid by taxpayers but cannot measure the anxiety, frustration, and anger associated with filing." Joseph Bankman, *Simple Filing for Average Citizens: The California Ready Return*, in Tax Notes 1431, 1431 (June 13, 2005).

53. Adam Smith wrote:

A tax may either take out or keep out of the pockets of the people a great deal more than it brings into the public treasury, in the four following ways . . . Fourthly, by subjecting the people to the frequent visits, and the odious examination of the tax-gatherers, it may expose them to much unnecessary trouble, vexation, and oppression; and though vexation is not strictly speaking, expense, it is certainly equivalent to the expense at which every man would be willing to redeem himself from it.

See SANDFORD ET AL., *supra* note 35, at 24-25 (quoting ADAM SMITH, *Of Taxes*, in INQUIRY INTO THE NATURE AND CAUSES OF THE WEALTH OF NATIONS bk. 5, ch. 2, pt. 2 (n.p. 1776)).

54. Jeff Pope refers to an Australian and a Spanish study as currently, the two most advanced in this area. Jeff Pope, *Tax Compliance Cost*, in MARGARET LAMB, ANDREW LYMER, JUDITH FREDMAN, & SIMON JAMES eds., TAXATION – AN INTERDISCIPLINARY APPROACH TO RESEARCH (2005), 211 [hereinafter *Tax Compliance Cost*]. However, these studies' limitations show that the psychological-cost literature has not flourished. For the Australian study, see Robin Woellner et al., *Taxation or Vexation – Measuring the Psychological Costs of Tax Compliance*, in TAX'N COMPLIANCE COSTS 35, 44-46 (2001) (examining how a new statutory drafting style affected psychological cost, and concluding that while the pilot study highlighted a number of important factors to consider and correct, further study would be valuable); for the Spanish study, see Consuelo Diaz & Maria Luisa Delgado, *Personal Income Tax Compliance Costs in Spain*, in TAX COMPLIANCE COSTS MEASUREMENT AND POLICY 210, 220-21 (Cedric Sandford ed., 1995) (surveying taxpayers, but using questions that elicited responses concerning attitudes towards tax obligations in general rather than responses illuminating the psychological cost to users of performing tasks for their tax affairs).

frustration in dealing with their tax affairs; some employ a professional adviser primarily to reduce this burden of worry. In so far as this has the desired effect, the psychic cost then becomes a monetary cost.⁵⁵

The scholarly tax literature demonstrates a continued interest, at least in principle, in elucidating aspects of taxpayers' performance that are labeled "psychological cost."⁵⁶ Jeff Pope observes that psychological cost is an "important area [that] has been theoretically recognized in nearly all studies since Adam Smith . . . but then generally consigned to the 'too hard basket.'"⁵⁷ If actually incorporated into inquiries, the concept of psychological cost could encompass occurrences such as stress, anxiety, and frustration. How these occurrences should be measured and comprehended as psychological cost, however, has not been resolved in practice, as the desire to estimate psychological occurrences in monetary terms is as prevalent as the corresponding disappointment that this goal has not proven feasible.⁵⁸

3. International Standard Cost Model

The International Standard Cost Model ("SCM") is another compliance-cost concept, which is particularly well developed.⁵⁹ The SCM is, according to OECD, the most widely-applied methodology for measuring the costs of businesses' compliance with legal obligations and attainment of entitlements, and is

55. SANDFORD ET AL., *supra* note 35, at 18. The examples and surveys to which they subsequently refer to involve groups of taxpayers, mostly elderly, who experience difficulties in fulfilling their tax obligations and consequently experience what the authors call psychic cost. *See id.* at 195-96 (quoting extensively from CEDRIC SANDFORD & ALAN LEWIS, *The Poor Have Tax Problems Too*, ACCOUNTANCY 94, 95 (April 1986)). Note, however, that those studies do not attempt to measure psychic cost directly. *Id.* The studies' conclusions are instead based almost entirely on an assumed correlation between users' difficulties in figuring out how to perform their tax affairs, and the presence of aspects deemed to be psychic cost. *Id.* This approach corresponds to the authors' belief that psychic cost is difficult or impossible to measure. *Id.*

56. *See generally Tax Compliance Cost, supra* note 54, at 211.

57. Pope, *Tax Compliance Cost, supra* note 54, at 211.

58. *See* Woellner et al., *supra* note 54, at 46 ("It may even be possible to measure the psychological costs in some empirical way [a]lthough further consideration has still to be given to the determination of a monetary value that can be attributed to such cost and variations between them . . ."); DIAZ & DELGADO, *supra* note 54, at 221 ("The total compliance cost - leaving apart the psychological cost incapable of being evaluated in monetary terms . . .").

59. The SCM approach applies the designation "administrative costs" for the aspects of compliance cost it includes. SCM MANUAL, *supra* note 30, at 6-7.

used in most European countries.⁶⁰ The SCM approach achieves a high degree of detail by breaking down legal authorities, referred to as “regulation,” into manageable, measurable components.⁶¹

The SCM approach does this through two types of dissection: first, division of legal authorities into “information obligations”⁶² and “data requirements”⁶³, which constitute the potential bases for individual assessments; and second, division of taxpayer performance into discrete “activities.”⁶⁴

The structure of the SCM method is illustrated in Figure 1.⁶⁵

60. FORUM ON TAX ADMINISTRATION, *supra* note 34, at 14.

61. SCM MANUAL, *supra* note 30, at 8.

62. Information obligations are defined as follows:

Information obligations (IO) are the obligations arising from regulation to provide information and data to the public sector or third parties. An IO does not necessarily mean that information has to be transferred to the public authority or private persons, but may include a duty to have information available for inspection or supply on request. A regulation may contain many information obligations.

SCM MANUAL, *supra* note 30, at 8.

63. Data requirements are defined as follows:

“Each information obligation consists of one or more data requirements. A data requirement is each element of information that must be provided in complying with an IO.” *Id.* at 8. “Examples of typical data requirements: An information obligation may lay down requirements with regard to the following information: Identity of business – name and business registry number, etc.; Business’s turnover/statement of turnover; Statements of business’s equity.” *Id.* at 25, box 12.

64. The SCM Manual defines “activities” (which it designates as “administrative activities”) as follows:

To provide the information for each data requirement a number of specific administrative activities must be undertaken. The SCM estimates the costs of completing each activity. Activities may be done internally or be outsourced (i.e. done externally). It may be necessary to make acquisitions to complete specific activity and where these are only used in complying with the requirement they are included in the estimate.

Id. at 8.

65. *Id.* at 9.

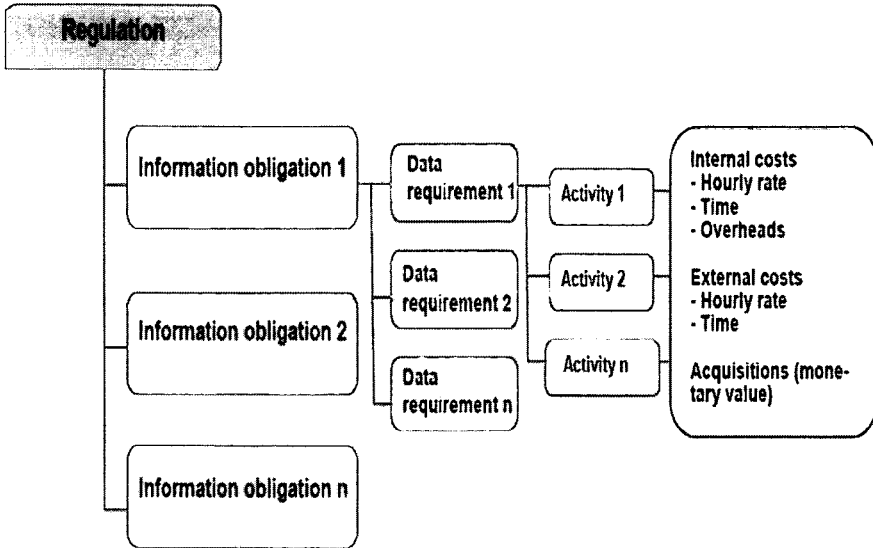


Figure 1

The list below, adapted from the SCM Manual, describes various “activities” in which users frequently engage in the income tax context. It is the undesirable occurrences deriving from these activities, such as time spent on the work of calculation and record-keeping, which constitute compliance cost under the SCM approach to the assessment of what taxpayers encounter.⁶⁶

- Familiarization: the user’s work of familiarizing herself or himself with the legal authorities for a given task, such as obtaining a charitable contribution deduction.
- Information retrieval: retrieval of the relevant figures and information needed to comply with a given obligation or right, such as finding the statements from a charitable association evidencing a paid contribution.
- Assessment: determination of which figures and information are necessary for the revenue body to accept the filing.
- Calculation: performance of the calculations necessary for the revenue body to accept the filing.
- Presentation of figures: presentation of the calculated figures, in forms or otherwise.

66. *Id.* at 5 (“The SCM methodology is an activity-based measurement . . .”).

- Checking: checking of the calculated figures, such as by reconciliation with other data.
- Settlement/payment: submission of tax payment.
- External meetings: meetings held with a tax preparer, lawyer, or the like.
- Copying, distribution, recordkeeping, etc.: maintenance of tax records.
- Reporting, submitting information: submission of information to the revenue body, including documentation, such as by filing a tax return.⁶⁷

These kinds of specific activities, which users perform to carry out tax-related tasks, are the source of various undesirable empirical occurrences, such as consumption of time and creation of stress.⁶⁸ However, the SCM approach, as a true compliance-cost concept, assesses only occurrences related to economic cost, such as time and cost of acquisition.⁶⁹

B. *Notions of Tax Complexity*

This Section describes some of the current notions of tax complexity that appear in the tax literature.⁷⁰ The descriptions aim at elucidating to what extent these complexity notions might be suitable for analyzing what taxpayers encounter when they perform activities like those listed above, and to analyze these

67. See *id.* at 25-26. Concerning the activity of familiarization, see also discussion *infra* note 91.

68. For each administrative activity, the SCM approach collects a number of cost parameters for the purpose of estimating cost per activity:

Price: Price consists of a tariff, wage cost plus overhead for administrative activities done internally, or hourly cost for external service providers.

Time, the amount of time required to complete the administrative activity.

Quantity: Quantity comprises of the size of the population of businesses affected and the frequency that the activity must be completed each year.

Combining these elements gives the basic SCM formula: Cost per administrative activity (or per data requirement) = Price x Time x Quantity (population x frequency).

Id. at 8-9 (emphasis deleted).

In addition, certain acquisitions may be included as an average cost per year, based on how long the purchased item is expected to last. See *id.*

69. See discussion *supra* note 68. The SCM approach has been criticized for this limitation: "It is well-known that SCM does not always adequately account for the experienced burdens on business. Obligations, which might not be very costly according to the SCM can be very annoying to business – and vice versa." INV. CLIMATE ADVISORY SERVS., WORLD BANK GRP., REVIEW OF DEN.'S PROGRAM FOR BETTER BUS. REGULATION 6 (2009).

70. For more notions of tax complexity and tax simplicity, including notions that do not pertain to what occurs empirically for taxpayers, see, e.g., Cooper, *supra* note 2, at 242.

occurrences in a way that provides adequate guidance to policymakers developing initiatives to ameliorate undesirable occurrences. Follow-up discussions regarding these notions appear in Parts III and V.

1. Diverse Notions of Tax Complexity

One notion of tax complexity is offered by Slemrod, who uses “complexity” as a stand-in for “cost”, more specifically “total resource cost”.⁷¹ Consequently, if tax complexity were estimated as the cost to taxpayers of managing their tax affairs, then this notion of tax complexity would be a synonym for taxpayers’ compliance cost.⁷² Another idea of tax complexity is presented by Louis Kaplow in his description of what constitutes more complex rules, pertaining specifically to the degree of differentiation in legal authorities:

A tax system may be complex and thus involve greater compliance costs for many reasons and in many ways. For present purposes, I shall focus on one important source of compliance cost: those arising from the need to make the tax base more accurately reflect taxpayers’ relative taxpaying ability. Such compliance cost may arise on account of more complex rules, for example, rules that attempt to distinguish more precisely between deductible business expenses and nondeductible personal expenses, or rules that include a range of fringe benefits in income. . . . It should be emphasized that the immediate costs of complexity cannot be measured by looking solely at the rules themselves. The primary source of compliance cost involves taxpayers’ behavior, often involving recordkeeping, so estimates of incremental

71. See *supra* Part II.A.1.

72. *Optimal Tax Simplification*, *supra* note 43, at 159. Slemrod’s definition includes resource-spending by the tax authorities, *id.* at 159, but this inclusion does not affect the use of “complexity” as a synonym for “cost.” See *id.* Slemrod sometimes explicitly refers to complexity as “cost.” See *id.* (“[I]t has become more complicated (or, specifically, more costly) to collect the revenue.”). Likewise, Gale and Holtzblatt, in their application of Slemrod’s definition, sometimes emphasize that “complexity” is equivalent to cost in their terminology, writing, “The complexity, or total resource cost, of the current system can be divided into several components . . .” William G. Gale & Janet Holtzblatt, *Role of Administrative Issues in Tax Reform: Simplicity, Compliance, and Administration*, in UNITED STATES TAX REFORM IN THE TWENTY-FIRST CENTURY 185 (2002).

paperwork costs are usually more probative than counts of number of pages of rules.⁷³

Kaplow's explanation implies a distinction between descriptions of legal authorities (tax rules), on the one hand, and assessments of what happens for users, on the other.⁷⁴ He assesses the latter as compliance cost.⁷⁵

A different use of the term "complexity" in the scholarly tax literature pertains to the "plain English" discussions, which emphasize linguistic issues.⁷⁶ Complexity here refers to the comprehensibility of legal authorities.⁷⁷ Bobbie Martindale states it in this way: "Two factors affecting comprehensibility are content complexity (the inherent difficulty of the subject matter) and text complexity (the readability of the writing and presentation style)."⁷⁸

An additional notion of "tax complexity" in the literature concerns situations in which the occurrences deriving from a tax-related task are appraised as reflecting an insufficient level of user performance.⁷⁹ According to this approach, tax simplicity or complexity consists of the level of optimality in the empirical occurrences.⁸⁰ If what the user encounters is appraised as poor, then the outcome of the performance warrants the label "complex", according to this notion.⁸¹ Complexity is thus the opposite of the characteristic of a performance whose empirical occurrences are appraised as satisfactory.⁸² An example of this designation appears in the work of Sidney Roberts, who specifies the user to whom his definition applies:

To the tax advisor, complexity means:

- 1) A reasonably certain conclusion, in some instances, cannot be determined despite diligent and expert research; or

73. Louis Kaplow, *How Tax Complexity and Enforcement Affect the Equity and Efficiency of the Income Tax*, 49 NAT'L TAX J. 135, 138 (March 1996).

74. *See id.*

75. *See id.*

76. *See Cooper, supra note 2*, at 239 (finding the "plain English" literature to be a sub-specialty in the field of taxation, but critical regarding its merit).

77. *See infra note 78 and accompanying text.*

78. Bobbie Cook Martindale et al., *Tax Law Complexity: The Impact of Style*, 29:4 J. BUS. COMM. 383, 383 (1992).

79. A good example of this designation comes from Sidney Roberts, as discussed immediately below. *See Roberts, supra note 2*, at 34.

80. *See id.*

81. *See generally id.*

82. *See generally id.*

- 2) A reasonably certain conclusion can be determined only after an expenditure that is excessive in time and dollars.⁸³

One more allusion to tax complexity, which bears mentioning, pertains to concerns about unfounded disparities in a tax regime.⁸⁴ In the terminology of complexity and simplification, the idea is that simplification should eradicate what are seen as complexity-engendering disparities.⁸⁵ Note, however, that the problem of disparities does not pertain solely to occurrences, like time consumption, that result from taxpayers' actual performance of compliance activities.⁸⁶ Rather, it pertains just as much to efficiency and equity matters, the two other grounds for evaluating tax policy.⁸⁷

2. Compliance Complexity, Transactional Complexity, and Rule Complexity

A common distinction among three kinds of complexity – compliance complexity, transactional complexity, and rule complexity – originates in David Bradford's *Untangling the Income Tax*.⁸⁸ Bradford writes:

83. Roberts, *supra* note 2, at 34. Cooper also recognizes this as one of the conceptions of tax complexity and simplification present in the literature: "The rule chosen would not be simple if it is difficult and excessively costly for taxpayers to comply with." Cooper, *supra* note 2, at 242.

84. See *Current Federal Income Tax and the Need for Reform: Hearing Before the H. Comm. On Ways and Means*, 112th Cong. 3 (2011) (statement of Martin A. Sullivan, Economist and Contributing Editor, Tax Analysts) ("The essence of an efficient and competitive system is a level playing field.").

85. With regard to this idea, Sullivan offers the following:

A quarter of a century ago, President Reagan defied all the skeptics and provided the leaderships for a bipartisan overhaul of the tax system that dramatically lowered tax rates and scaled back tax breaks that favored certain groups and activities over others. It was a victory over the special interests. Twenty-five years later the need for tax reform is greater than ever.

Id. at 1.

86. See *id.* ("The perception of unfairness . . . is an insult to the majority of taxpayers bearing their fair share. And on top of all this our tax code is dead weight on the shoulders of the American economy.").

87. See *supra* Introduction and note 1.

88. DAVID F. BRADFORD, *UNTANGLING THE INCOME TAX*, 266-67 (1986). See also Steven A. Dean, *Attractive Complexity: Tax Deregulation, The Check-The-Box Election, and the Future of Tax Simplification*, 34 HOFSTRA L. REV. 405, 418 (2005-06) (using Bradford's categories, such as compliance complexity); McCaffery, *supra* note 2, at 1270-71 n.18 (stating that rule complexity, which he terms "technical complexity", is the "first basic understanding of simplification"); Paul, *supra* note 5, at 154 n.4 (offering a complexity framework that builds on Bradford's framework, among others); William J. Gale & Jeffery Rohaly, *Effects of Tax Simplification Options*, in *THE CRISIS IN TAX ADMINISTRATION*, 303, 307 (Henry J. Aaron & Joel Slemrod eds., 2004) ("Scholars often

Simplicity in taxation has various aspects, and often a change that simplifies in one way introduces greater complexity in another. We may distinguish three kinds of complexity: compliance complexity (referring to the problems faced by the taxpayer in keeping records, choosing forms, making necessary calculations and so on); transactional complexity (referring to the problems faced by taxpayers in organizing their affairs so as to minimize their taxes within the framework of the rules); and rule complexity (referring to the problems of interpreting the written and unwritten rules). These three forms of complexity are related, but improving the law with respect to one form may make it worse with respect to the others.⁸⁹

The distinguishing features listed by Bradford appear to correspond to characteristics of taxpayers' various activities, such as tax planning or activities related to filing.⁹⁰ Thus, rule complexity can be understood as referring to the activity of familiarization with obligations and entitlements pursuant to the relevant legal authorities. Compliance complexity can be understood as referring to activities such as computation, filing, and payment.⁹¹ Transactional complexity can be understood as referring to certain activities beyond those immediately necessary to fulfill compulsory tasks. This can be the case, for example, when users have the opportunity to use tax planning to utilize disparities in a tax regime. Such optional tasks would involve additional activities performed by taxpayers, such as constructing different possible scenarios for taxable income,

speak of three kinds of complexity – rule complexity, transactions complexity, and compliance complexity.”).

89. BRADFORD, *supra* note 88, at 266-67.

90. *See id.*

91. Compliance complexity can thus be understood as designating activities such as those the International Standard Cost Model describes as common “activities”, with the exception of the activity of familiarization. *See supra* Part II.A.3. The SCM approach also differentiates between familiarization and the other sorts of activities. *See* SCM MANUAL, *supra* note 30, at 25. (“As regards the first administrative activity, the resources used by business to *familiarize themselves with the information obligation* [legal authority], it is linked to the actual information obligation, while the rest of the administrative activities are linked to the [fulfillment of the] individual data requirement.”). The activity of familiarization appears under Bradford’s designations, as discussed above, to be embraced by “rule complexity.”

See also infra note 110 (discussing the activity of familiarization with legal authorities).

comparing them, and choosing the most favorable one. Notably, in a situation involving undertakings designated as transactional complexity, taxpayers most likely would still have to carry out the activities of familiarization as well as calculation, reporting, and the like – that is, the activities to which rule complexity and compliance complexity refer. Transactional complexity under this reading of Bradford refers mainly to some specific additional activities.

Under this interpretation, any situation in which a taxpayer undertakes a recordkeeping activity would constitute compliance complexity, and likewise, any optional planning activity would constitute transactional complexity. Hence, the sub-terms of complexity, such as “compliance complexity,” are used to designate particular activities, but they do not designate or distinguish between different types of empirical occurrences that derive from performance of the various activities.

Another interpretation of Bradford’s set of distinctions would be that within what are designated “compliance”, “transaction”, and “rule” complexity, an instance would be more or less complex based on how problematic it is. This would be a reading à la Roberts’ use of the term “complexity.”⁹² Under such an interpretation, the question would arise as to what distinguishes a problematic activity from a non-problematic or less problematic one. When, for instance, does keeping records, choosing forms, or making necessary calculations become a problematic activity? In other words, what occurs in carrying out such activities that constitutes complexity, and how should it be assessed?

A third interpretation of Bradford’s notion of complexity would be that it simply pertains to compliance cost, and consequently that all activities constitute complexity. This would correspond to Slemrod’s notion of complexity as it pertains to taxpayers’ cost.⁹³ Parts III and V expound on the competence of the aforementioned tax complexity notions to assess what users encounter in performing tax-related tasks.

III. CONFIGURATION OF THE CONCEPTS FOR ANALYSIS IN THE FIELD OF TAX COMPLEXITY

The term “tax complexity” is currently used in a confusing variety of ways.⁹⁴ Its meaning is fairly clear when it is used to

92. See *supra* Part II.B.1.

93. See *supra* Part II.A.1.

94. See *generally supra* Part II.

designate a field of tax policy.⁹⁵ Aside from that well-understood accepted usage, however, when the term “tax complexity” is used to refer to specific phenomena that occur in the administration of tax matters, this Article recommends that it should be clearly explicated with reference to whatever concrete phenomena it is meant to include.⁹⁶ Otherwise, there is a great risk of misunderstanding caused by numerous distinct phenomena being called the same thing, and by the haziness of some of the complexity notions.⁹⁷ Currently, these two factors – the semantic uncertainty and the lack of clarity in some of the notions being referred to – work to impede effective analysis concerning taxpayers’ compliance work, and consequently have a negative effect on the gathering of information vital for policymakers’ work toward improvement.⁹⁸

Moreover, there is an additional reason the term “tax complexity” is regarded as elusive when pertaining to taxpayers’ empirical experiences. This Article argues that the tax literature has underestimated the challenges involved in simply assessing what taxpayers encounter, as an empirical matter, in handling their tax affairs.⁹⁹ An assessment of empirical occurrences deriving from performance of a task is essentially a study of human-product interaction or human-task dealing.¹⁰⁰ People are a diverse lot, and situations vary; as a consequence, assessing outcomes of human behavior is inherently difficult.¹⁰¹

In response to these issues, this Part presents a configuration of the concepts and tax complexity notions reviewed in Part II *supra* with the goal of clarifying these concepts, and how they relate to each other, by identifying the diverse underlying issues which each of them encompasses. The configuration is based on the usability *framework*. The usability literature is intimately concerned with the challenges of analyzing and assessing human task performance in specific

95. See *supra* note 3 and accompanying text.

96. See *infra* Part V.A.

97. See *supra* Part II.B.

98. See *generally supra* Part II.B.

99. See *infra* Part III.C.4.

100. See *e.g.* SHNEIDERMAN & PLAISANT, *supra* note 12, at 12, making a like point with regard to human performance with human-computer interaction: “The interdisciplinary design science of human-computer interaction began by combining the data-gathering methods and intellectual framework of experimental psychology with the powerful and widely used tools developed from computer science.”

101. See *id.* at 25 (noting the challenges of physical, cognitive, perceptual, personality, and cultural differences).

contexts,¹⁰² and this Article demonstrates how the ISO usability framework embeds principles that generally can assist in obtaining validity in assessments of empirical occurrences derived from compliance-task performance.

The usability framework applies the term “outcome of use” for the empirical occurrences as a whole,¹⁰³ and this Article adopts this use of the term. Hence, the term “outcome of use” encompasses what actually occurs for users in their management of tax-related tasks, and the diverse specific occurrences can be evaluated in different ways depending on what concept is applied for assessment, such as the compliance-cost concept or the usability concept.¹⁰⁴

The discussions in this Part also serve as a background for Part IV’s presentation of the usability *concept*, which augments the types and comprehension of occurrences that may be assessed.

A. *Context of Use*

One of the usability framework’s main terms is “context of use,”¹⁰⁵ which this Article adopts, along with “outcome of use.” All activities happen in a context – a real-world setting in which users are actually performing tasks.¹⁰⁶ For an assessment of empirical occurrences to truly elucidate taxpayers’ experiences, and thus to be valid, the assessment setup needs to include a depiction which sufficiently reflects the users’ actual settings.¹⁰⁷ Obviously, the actual settings in which different user segments handle their tax affairs could be described in various ways. This Section elaborates on the depictions of real-world settings – contexts of use – offered by the ISO usability framework. Figure

102. See ISO 9241-11, *supra* note 12, at iv (explaining that the objective for usability evaluation is to enable users to achieve goals and meet needs in a particular context of use). See also INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, ERGONOMIC REQUIREMENTS FOR OFFICE WORK WITH VISUAL DISPLAY TERMINALS (VDTs) - PART 1: GENERAL INTRODUCTION, ISO 9241-1 (1997) (“ISO 9241 emphasizes the need to specify the factors affecting the performance of the users, and the need to adopt a user-performance approach to evaluate systems.”) (the general introduction to the ISO 9241 standard).

103. ISO 9241-11, *supra* note 12, § 5.1.2, Figure 1.

104. The tax liability itself, and how it affects the taxpayer with regard to allocation effect, is not part of compliance cost, see *supra* note 35 and accompanying text, or an aspect assessed as part of outcome of use.

105. See ISO 9241-11, *supra* note 12, § 3.5.

106. *Id.*

107. See ISO 9241-11, *supra* note 12, at Annex B, B.4; see also *infra* Part III.B. (explaining that outcome of use is not intrinsic to the legal authority inducing the task, but is influenced by the specific setting of the actual performance).

2 below shows how ISO identifies the components of context of use.

As another part of the setup for an assessment, the particular performance, whose empirical occurrences are to be assessed, must be identified.¹⁰⁸ This is designated in Figure 2 below as the “basis” of the assessment. For example, a basis of assessment could be the legal authority authorizing deductions for charitable contributions. In that case, the outcome of use to be assessed will consist of occurrences deriving from the user segment’s performance of the activities necessary to obtain the deduction. This form of assessment setup corresponds to the methodology of the International Standard Cost Model.¹⁰⁹

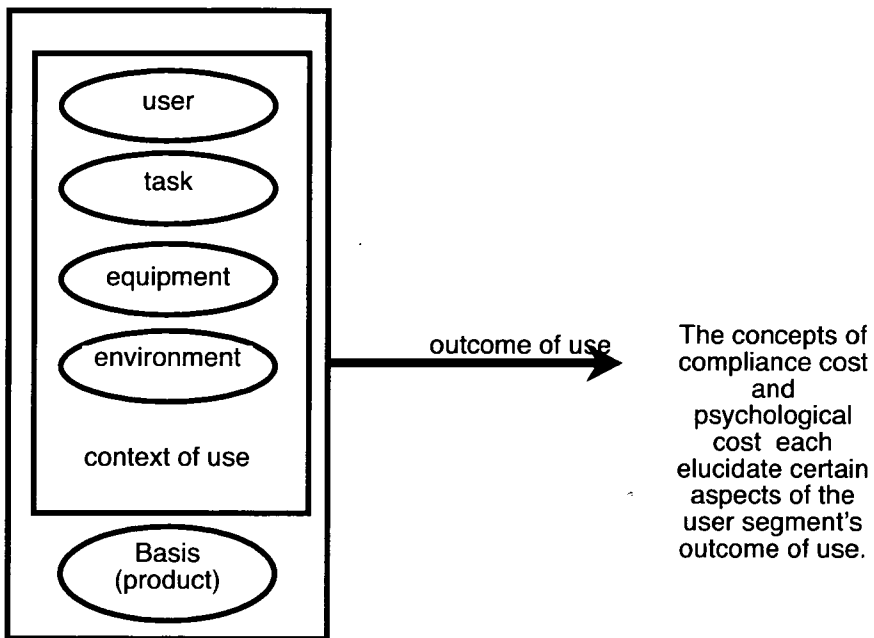


Figure 2¹¹⁰

108. See ISO 9241-11, *supra* note 12, §§ 3.10 & 7.1.

109. See *supra* Part II.A.3. (presenting “regulation,” “information obligation,” and “data requirement” as increasing levels of detail within the “basis”).

110. See ISO 9241-11, *supra* note 12, § 5.1.3. The Figure is taken from ISO, but has been altered here to designate a product as “basis” and to omit the right side of the

As mentioned, to ensure a valid assessment, the depiction of the setting of performance must cover enough characteristics of the user's actual setting.¹¹¹ The subsections that follow describe the components of context of use proposed by the usability framework to meet this requirement: users, tasks, and equipment/environment.

1. Users

The abilities of the person, who performs the activities required for a task, might significantly influence the outcome of use, which makes the depiction of those abilities important.¹¹² ISO describes the depiction of users required for an assessment in this way:

Relevant characteristics of the users need to be described. These can include knowledge, skill, experience, education, training, physical attributes, and motor and sensory capabilities. It may be necessary to define the characteristics of different types of user, for example users having different levels of experience or performing different roles.¹¹³

A practicable approach for depicting users' abilities could be accomplished by assigning them to different user segments, based on the most important segregating characteristics. For individual taxation, one example of a notable difference in ability is found between responsible taxpayers, managing the tasks by

original ISO figure, which refers specifically to usability measures. *See also infra* Figure 5 and accompanying text.

It may be helpful to think of legal authorities as having a dual role: they are both normative texts and media for promulgation of this normative matter. In legal authorities' capacity as normative texts, constituting obligations and entitlements, they are not used for the performance in a direct sense. Legal authorities in this capacity are better understood as imposing tasks, such as the work required to claim a charitable-contribution deduction. Since legal authorities in this capacity are not "products" that are used in a literal sense, this Article instead employs the broader term "basis." When used as media for promulgation of the normative matter, legal authorities can be understood literally as products, and in that sense are similar to products such as tax software programs. In their capacity as media for promulgation, legal authorities can be related to how well taxpayers can perform the activity of familiarization, and thus understand the normative matter, when they do so directly, by reading the legal authority itself. Evidently, however, many taxpayers' familiarization activity does not occur with the help of the legal authority itself, but instead through the use of derivative texts such as the IRS's instructions, privately-published tax guides, computer programs, and the like.

111. *See* ISO 9241-11, *supra* note 12, at Annex A.

112. *See* ISO 9241-1, *supra* note 102, at iv (emphasizing the importance of taking variation in user skills into account).

113. ISO 9241-11, *supra* note 12, § 5.3.1.

themselves as laymen, and professionals, from whom taxpayers purchase assistance.¹¹⁴

2. Tasks

The “tasks” component encompasses the taxpayer’s broader compliance goals, as she or he intends to fulfill them.¹¹⁵ Examples of compliance tasks are completing (or attempting to complete) a specific information-reporting obligation or particular deduction, or, even more comprehensively, filing an entire tax return. “Activities”, by contrast, are the types of work that are necessary to execute tasks.¹¹⁶ The SCM’s list of commonly-performed administrative activities – for example, information retrieval, calculation, and submission of payment¹¹⁷ – illustrates activities that are involved in executing broader tax-compliance tasks.¹¹⁸ Outsourcing, as when an individual taxpayer purchases assistance, can influence which activities a segment of taxpayers might do themselves, such as collecting and organizing bills for potential deductible costs before submitting them to the tax preparer.¹¹⁹

The SCM’s list of commonly-performed activities is compiled with a view toward compulsory tasks and common voluntary tasks.¹²⁰ However, other kinds of tasks might lead users to perform additional types of activities. For example, a taxpayer might do sophisticated tax planning, and that task might involve additional activities such as comparing possible scenarios for taxable income. Thus, the obligations and entitlements set forth in the legal authority, which is the basis for the assessment might be handled in different ways involving different activities. In that case, various assessments might be necessary to ensure

114. See, e.g., Robert Kidder & Craig McEwen, *Taxpaying Behavior in Social Context: A Tentative Typology of Tax Compliance and Noncompliance*, in 2 TAXPAYER COMPLIANCE 47, 58 (Jeffrey A. Roth & John T. Scholz eds., 1989) (stating that complying with tax-law procedures is a “luxury” because it requires “skill and resources that relatively few people have”); Deborah H. Schenk, *Simplification for Individual Taxpayers: Problems and Proposals*, 45 TAX L. REV. 121, 128 (1989-90) (suggesting that, “[f]or many taxpayers, the tax return and instructions present a bewildering morass of rules which cannot be mastered easily”).

115. See ISO 9241-11, *supra* note 12, § 3.9.

116. See ISO 9241-11, *supra* note 12, § 5.3.2.

117. See *supra* Part II.A.3.

118. See SCM MANUAL, *supra* note 30, at 20-26.

119. See generally SCM MANUAL, *supra* note 30, at 30, 32, 35.

120. SCM MANUAL, *supra* note 30, at 16-17 (discussing whether to measure voluntary tasks considered necessary to follow (because the majority of the business for which the rule is relevant choose to observe the rule) and tasks not regarded as necessary).

sufficiently accurate portrayal of the actual spectrum of activities performed by different user segments.¹²¹

3. Interplay with Environment

The performance of activities to advance tax-related tasks, like the performance of activities to advance any task, takes place in a physical and social environment. The usability framework depicts this through two components, which it terms "equipment" and "environment."¹²² The "equipment" component depicts that with which the performance is integrated or interacts directly, whereas the "environment" component

121. ISO 9241-11, *supra* note 12, § 5.5. ISO emphasizes that any description of the activities and steps involved in performing the tasks will be related to what the user is attempting to achieve. This is important in a tax context when there are options for tax planning, since these will cause some but not all users to perform additional activities in order to achieve tax planning. The ISO's description is as follows (ISO refers to usability, but the principle is the same for other assessments of outcome of use such as compliance cost): "For a general-purpose product, it will generally be necessary to specify or measure usability in several different representative contexts, which will be a subset of the possible contexts and of the tasks which can be performed. There may be differences between usability in these contexts." ISO 9241-11, *supra* note 12, § 5.5.

The choice of what to assess among the different tasks possible under a legal authority depends on the purpose of the inquiry. For example, the SCM Manual discusses the question whether to measure compliance cost incurred from voluntary tasks, as opposed to measuring compliance cost only from compulsory tasks, and leaves this to the researcher to decide, based on the specific research interest. See SCM MANUAL, *supra* note 30, at 16. The SCM Manual's distinctions between compulsory and voluntary tasks are in accordance with earlier literature on the topic. Writing under the heading "The Cost of Tax Planning," Sandford et al. elaborate on this issue with regard to business. They explain:

If we consider business costs, it is sometimes possible to separate tax planning from the purely computational aspects of tax compliance, but such separation has little value. The time of investment, or arranging cash flows to maximize benefit from tax provisions, are the kind of measures which we would expect from "a reasonable man" . . . With big commercial transactions, such as mergers, or the transfer between countries of assets of a multi-national company, there is, of necessity, a large input of tax planning. In the literal sense it is avoidable, but any company, which ignored it could find itself in serious trouble. Such tax planning, which consists of a detailed examination of the implications of a transaction and the choice of a method, which minimizes tax, is an essential cost of compliance, a necessary ingredient of commercial activity.

Sandford et al., *supra* note 35, at 12-13. The kinds of activities Sandford describes are what are often designated as "transactional complexity." See *supra* Part II.B.2. The quotation helps to elucidate how the outcome of use from performing activities for such tasks can be analyzed similarly to user performance for other tasks. The particularity of voluntary tax-planning tasks is thus not that they require a different mode of analysis as such, but rather that it must be determined whether policymakers' goals for an inquiry include evaluation of taxpayers' performance of those kinds of tasks.

122. See ISO 9241-11, *supra* note 12, § 5.3. To which of these two components a particular piece is assigned can be less important than its inclusion in the depiction of context of use.

represents the external setting.¹²³ Examples of portions of actual settings that would be included in the “equipment” component could include a required special form for charitable deductions or tax software utilized by taxpayers.

ISO describes “environment” as including the “wider technical environment (e.g. the local area network), the physical environment (e.g. workplace, furniture),” and the “social and cultural environment (e.g. work practices, organizational structure and attitudes).”¹²⁴ Relevant examples in the income-tax context could be computer hardware, GPS systems (as they relate to mileage deductions), smartphones (for on-the-spot electronic recordkeeping), and the like. Another example of the “environment” component could be the form in which financial data are obtained. Some users might have to compile income from paper receipts, while others might receive one concise statement of income.

The purpose of the components of the context of use, taken together, is to depict the actual setting, and the formal distinction between equipment and environment merely offers a structure for that depiction.¹²⁵ Overall, ISO emphasizes that the depiction of context of use should be detailed enough to include the circumstances that have considerable influence on the outcome of use¹²⁶ for tax-compliance tasks, for instance, this could be the use of tax-preparation software or prefilled tax returns.¹²⁷

123. See ISO 9241-11, *supra* note 12, §§ 5.3.3, 5.3.4.

124. ISO 9241-11, *supra* note 12, § 5.3.4.

125. See, e.g., Abran et al., *supra* note 20, at 329 (offering alternative explanation of context characteristics in the ISO framework by simply applying the three components: users, tasks, and environment).

126. See ISO 9241-11, *supra* note 12, at § 5.1.3.

Attempts to specify to some extent the real-world setting for an assessment of “complexity” are not new in the field of tax complexity and simplification. However, the usability framework emphasizes the importance of making the portrayal of a setting sufficiently representative of the actual setting, and consequently the importance of making the portrayal adequately comprehensive. In the tax complexity literature, an example of a partial specification of the actual setting is Roberts’ complexity notion, *supra* Part II.B.1., which specifies that the user is a tax adviser. Another example is McCaffery, who states that there is no single comprehensive definition of simplicity and argues that “the task of ‘simplicity’ first requires us to face an intricate set of perspective-related threshold questions: simplification in what sense? Simplification to whom? Simplification to what extent?” McCaffery, *supra* note 2, at 1270. Very roughly stated, the first of McCaffery’s questions concerns what for the usability framework is the task component; the second question concerns the user component; and the third question concerns what the “basis” of the assessment will be. *Id.*

127. An example of a prefilled tax return system is “ReadyReturn,” a service of the State of California. See State of California Franchise Tax Board, *Your California Tax Return May Be Ready and Waiting For You*, <https://www.ftb.ca.gov/readyreturn/> (last visited Oct. 10, 2012) (“ReadyReturn is a free service we developed to make filing

B. *Outcome of Use is Not Intrinsic to a Legal Authority*

This Section argues for the importance of distinguishing between assessments of empirical occurrences and descriptions of legal authorities.

Portrayal of the real-world setting in the form of context of use is essential to a valid assessment because outcomes of use from task performance are heavily influenced by specific settings.¹²⁸ This is why occurrences, whether measured as compliance cost or usability level, cannot be evaluated merely by examining the basis of the assessment, such as the legal rule or authority that induces the obligation or allows the entitlement (in ISO terminology, the “product” that is being used). ISO explains the basic idea as follows:

[T]he attributes which a product requires for usability depend on the nature of the user, task and environment. A product has no intrinsic usability, only a capacity to be used in a particular context. Usability cannot be assessed by studying a product in isolation.¹²⁹

Consequently, as outcome of use is not an intrinsic phenomenon, any variable, such as the user’s skills,¹³⁰ may change the outcome of use that results from performing the activities for a tax-related task induced by a legal authority.

The principle that outcome of use is a concrete and context-dependent empirical phenomenon, distinguishable from a description of a static product or legal authority, is independent of whether a user’s encounter is measured as level of usability in the performance, as compliance cost, or as something else. The following example illustrates this, and shows how a compliance-

individual income tax returns easier. We use information the state already has from the last return you filed and from your Form W-2 to pre-fill a California state tax return. If you qualify, your return may be ready and waiting for you. All you need to do is review your return, make any necessary changes, and file your return.”). A survey of people who actually used ReadyReturn has shown it to be highly appreciated as easing the burden of tax compliance. See Bankman, *supra* note 52, at 1432. Internationally, pre-filing has evolved into a significant strategy for lightening the burden of tax compliance. See FORUM ON TAX ADMINISTRATION: TAXPAYER SERVICES SUB-GROUP, SURVEY OF TRENDS AND DEVELOPMENTS IN THE USE OF ELECTRONIC SERVICES FOR TAXPAYER SERVICE DELIVERY 78 (March 2010) (“there are a number of revenue bodies that now offer a fully automated ‘end-to-end’ process for return preparation, assessment and payment collection/refund crediting, thereby providing significant benefits to their taxpayers.”)

128. See the following discussion (referring to ISO 9241); see also ISO-9241-11, *supra* note 12.

129. ISO 9241-11, *supra* note 12, at Annex D, D.1.

130. See *supra* Part III.A.1.

cost study would assess actual occurrences in a specific context of use. Assume that a study shows that taxpayers who take a deduction for charitable contributions in a given year incur an average compliance cost of \$30 (the estimated value of their resource-spending) as a result of managing this deduction. Subsequently, some prominent celebrities participate in national television shows that recommend making charitable donations and explain how to carry out the activities to take the deduction. The next year, a study of the compliance cost for the charitable-contribution deduction shows that the average compliance cost has fallen to \$27, despite there having been no change in the law or administrative procedure. This example shows that compliance costs are not intrinsic to the legal authority. The attributes of the legal authority entitling taxpayers to charitable-contribution deductions have not changed, and thus cannot explain the different result. However, the context of use has changed, as users have become more knowledgeable and thus more capable, resulting in less resource-spending. This can explain the change in compliance cost. Thus, assessments of empirical occurrences that derive from performance of tasks induced by a legal authority are analytically distinct from descriptions of attributes of that legal authority, such as the legal authority's length or the number of distinctions or exceptions it contains.

Assessment of empirical occurrences can be characterized as *process-oriented*— it is taxpayers' multiplicity of encounters from carrying out activities for tax-related tasks that is being assessed. Process-oriented assessments concern empirical occurrences that imply welfare propositions. In contrast, a *product-oriented*¹³¹ description of a legal authority (i.e., a legal rule) — like a description of a physical product — is a static account of attributes of the object in question.¹³² Product-oriented analyses can articulate something about the features of

131. Concerning references to the terms process- and product-oriented, *see, e.g.*, Abran et al., *supra* note 20, at 326, for a discussion of the distinction between process-oriented and product-oriented assessments with regard to the various international standards developed by ISO.

132. A look in a dictionary shows that the term "complexity" can denote either something relating to the outcome of use, or a description of a product or other object. *See, e.g.*, WEBSTER'S II: NEW COLLEGE DICTIONARY 235 (3d ed. 2005). One definition, "Complex: difficult to understand due to intricacy," describes an outcome of use. Still, such a description leaves unanswered the question of exactly what constitutes "difficult," or, stated differently, what specifically happens to cause the outcome of use to be "complex." Webster's other definitions of complex include: (1) "Composed of interconnected or interwoven parts," and (2) "Composed of two or more units." These two meanings relate to attributes of an object.

legal rules and can make possible comparisons of variable characteristics and attributes of legal authorities, but such descriptions do not directly concern the occurrences, which taxpayers might encounter.¹³³

1. Description of Legal Authorities' Attributes

Outcomes of use and descriptions of legal authorities' attributes are, as established above, analytically distinct phenomena.¹³⁴ Consequently, product-oriented descriptions of legal authorities' attributes can stand independently.¹³⁵ If the tax code is the product to be described, then notable attributes of that product are its number of words.¹³⁶ Another salient product-oriented attribute is the number of pages required for a specific printing.¹³⁷ The possible relevance of such attributes to complexity research obviously depends on what one is examining.

A potentially more fruitful way to analyze attributes of legal authorities is to apply an abstract definition of complexity. A standard interdisciplinary definition is that something is complex if it consists of many elements and many possible relations.¹³⁸ By applying this or another general definition of complexity, it might be possible to estimate the level of complexity of a legal authority, such as the one governing

133. Kaplow's description of tax complexity and compliance cost, *see supra* Part II.B.1., embraces a distinction corresponding to what this Article presents as descriptions of legal authorities' attributes (product-oriented) and outcome of use (process-oriented).

134. Another way to make the same point is to note that outcome of use is not simply a property of a product – or of a legal authority – in isolation, but rather that outcome of use will also depend “on who is using the product, the goal that they are trying to achieve, and the environment in which the product is being used.” *See* JORDAN, *infra* note 201, at 7 (making the point in regard to usability).

135. In the terminology of usability, product attributes are also referred to as “design characteristics.” *See, e.g.*, JORDAN, *infra* note 201, at 25.

136. *See, e.g.*, National Taxpayer Advocate, 2008 Annual Report to Congress, IRS, Executive Summary 1 (“One count shows the number of words in the tax code has reached 3.7 million . . .”).

137. NTU's paper entails a substantial number of different phenomena referred to under the term “tax complexity.” *See* David Keating, National Taxpayers Union (NTU), *A Taxing Trend: The Rise in Complexity, Forms, and Paperwork Burdens*, 1 (April 15, 2009), <http://www.ntu.org/news-and-issues/taxes/income-tax/a-taxing-trend-the-rise-in-tax-complexity.html> (“If you downloaded the Code, pasted it into a Word document, and printed it, you better have plenty of paper and ink on hand because it would take 9,114 pages.”).

138. *See, e.g.*, NIKLAS LUHMANN, *EINFÜHRUNG IN DIE SYSTEMTHEORIE* 173 (Dirk Baecker ed., 2002); Various Authors, *Science About Complex Systems*, 284 *SCIENCE* 5411 (1999) (providing more elaborate definitions of complexity).

charitable deductions, based on the incidence of attributes that the definition regards as complex.¹³⁹

It may be the case that legal authorities for the income tax, in line with the general trend in society, are becoming more complex.¹⁴⁰ This could perhaps be ascertained by applying a complexity definition embracing attributes such as the number of articles, their interconnections, and the number of exceptions to main rules. For instance, a change made to the set of legal authorities governing charitable contribution deductions, making not all in-kind contributions eligible for a deduction, could serve as an example of rising tax complexity, given that such elaboration increases the number of statutory articles and the number of distinctions. However, increased tax complexity established in accordance with such a notion cannot be used to infer that users' outcome of use is deteriorating as a result.¹⁴¹

139. Deborah Paul offers one delineation of tax complexity, which she designates as "complication": "A 'complicated' regime, such as the federal income tax, consists of numerous detailed authorities". Paul, *supra* note 5, at 158. She explains that even though hard to count within legal authorities, "Nevertheless, large differences in complication are readily apparent. The federal income tax is clearly more complicated than the sales and use tax of the State of Wyoming." *Id.* at 159. What Paul designates as "complication" is, in the terminology of this Article, a product-oriented description, i.e., it describes attributes of legal authorities themselves. *Id.*

Paul proposes a model of tax complexity, which "distinguishes among three ways in which a tax regime may be complex. It may be "complicated," "intractable," or "incoherent." *Id.* at 157.

Tractability" concerns "the ease with which the regime's underlying concepts may be applied. Tractability makes the amount of a person's tax liability, as well as the time and manner of payment, as Adam Smith stated, 'clear and plain' to the taxpayer and 'every other person.' The federal income tax is intractable, relying on such difficult concepts as income, realization, dividend, and corporate business purpose.

Id. at 160. Despite Paul's references to difficult concepts, what she designates as "intractable/tractable" has great similarity to Roberts' notions of tax complexity, *supra* Part II.B.1., since it is the level of optimality in the assessed outcome of use which determines the labels complex/simple or, for Paul, intractable/tractable. Paul's third distinction, "incoherent/coherent," concerns the degree to which the purposes of the tax regime "are expressed in, and served by, the legal authorities. A coherent tax regime forms a logical whole . . . Incoherence generates social costs." *Id.* at 161. Whether the "incoherent/coherent" designation can be seen to relate to product-oriented descriptions or process-oriented assessments appears to depend on whether an ascertainment is supposed to be based solely on a static evaluation of the relation between purposes and legal authorities (a product description) or whether the empirical occurrences constituting level of social cost are the major determinant (process-oriented).

140. Regarding the rise in complexity in society generally, *see, e.g.*, ERIC D. BEINHOCKER, *THE ORIGIN OF WEALTH: EVOLUTION, COMPLEXITY, AND THE RADICAL REMAKING OF ECONOMICS* (2006); *see also generally* LARS QVORTRUP, *THE HYPERCOMPLEX SOCIETY* (2003).

141. *See supra* Part III.B. (establishing that descriptions of legal authorities' attributes are distinct from assessments of outcome of use).

Taxpayers who are subject to more complex rules, such as the rules in the above example limiting deductions to a certain range of in-kind contributions, could in fact have an outcome of use, which displays lower compliance cost or higher usability.¹⁴² In the example, this could be the case if the decrease in the number of deductible in-kind contributions results in segments of taxpayers having fewer tasks to perform, such as fewer deductions to take, or in their having fewer activities to execute, such as not having to estimate the value of an in-kind contribution.¹⁴³

2. Proxies and Causalities

Since outcome of use and the attributes of legal authorities are two logically separate phenomena, assessing one of them by estimating the other is an assessment through a proxy, which can be intricate and carries the risk of erroneous results.¹⁴⁴ For instance, the size of the tax code is an attribute of the product, whether assessed by number of sections, word count, or dimensions of the complete set of volumes. Attempting to use such a quantification of the tax code to assess aspects of use, such as level of achievement, time consumption, and stress, would be an application of a proxy, and would require an explanation as to why this proxy – the size of the code – would be appropriate for illuminating those occurrences.

Another consequence of the fact that outcomes of use and attributes of legal authorities are two distinct phenomena concerns the degree of certainty to which it can be established that specific attributes in fact cause particular effects on outcome of use.¹⁴⁵ Here it will suffice to mention only one prevalent topic from the field of tax complexity and simplification: the “plain English” discussion. The “plain English” literature purports to improve taxpayers’ outcome of use by redrafting tax laws, although only by changing their syntax in order to make the language clearer and more accessible.¹⁴⁶ Changing specific

142. For another example and discussion, see *infra* Part V.B.

143. Whether an appraisal of the imagined change should deem it an overall improvement is a separate question. Appraisal can be done in terms of compliance cost or usability. For different approaches to appraisal of assessed compliance cost, see *supra* Part III.A.2. For how one might appraise an assessed usability result, see *infra* Part IV.D.1.

144. See discussion *supra* Part III, Introduction (outcome of use) and *supra* Part III.B.1. (attributes of legal authorities).

145. See discussion *supra* Part III, Introduction (outcome of use) and *supra* Part III.B.1. (attributes of legal authorities).

146. See, e.g., HAYLEY ROGERS, *Drafting Legislation at the Tax Law Rewrite Project*, DRAFTING LEGISLATION: A MODERN APPROACH 77 (Constantin Stefanou & Helen

characteristics of statutory language affects attributes of legal authorities (i.e., is *product-oriented*).¹⁴⁷ However, as the goal of these changes is to improve outcomes of use, such as lowering resource-spending (i.e., is *process-oriented*), it is necessary to demonstrate that the attributes comprising “plain English” really lead to improvements in how well users carry out their tax-related tasks,¹⁴⁸ given their context of use. Obviously, such context of use would have to involve the relevant users actually consulting the legislation in question, and thus performing the activity of familiarization by reading and interpreting the legal authority itself.¹⁴⁹

C. *Configuration of Elements in Assessments, and Underlying Meanings of “Tax Complexity”*

Clarity as to exactly what issues the different kinds of analysis in the tax complexity field are concerned with is obviously important for policymakers’ work. To help avoid confusion, this Section offers a brief characterization of some of the many issues in the study of “tax complexity”.

This Article proposes that analysis related to users’ performance of tax-related tasks implies a primary determination as to whether a specific inquiry is to examine: (1)

Xanthakieds., 2008) (“The Tax Law Rewrite project . . . expresses its overall aim as, ‘to rewrite primary direct tax legislation to make it clearer and easier to use, without changing the law.’”).

147. See *supra* III.B.1.

148. See, e.g., Martindale et al., *supra* note 78, at 384 (“Increased comprehensibility could lead to more efficient tax return preparation and lower cost of compliance.”).

149. See *supra* note 110 (pointing out how many taxpayers’ familiarization activity does not occur with the help of the legal authorities themselves). Cf. Bobbie Cook Martindale et al., *supra* note 78, at 384, who argues confidently for causality between plain English and some improvement in outcome of use:

While readability of a law is not completely independent of its content, increasing the readability of United States tax law is feasible. Reducing text complexity will not cure the system’s problems, but it is a step in the right direction. Existing laws need only be recast in plain English.

Although few taxpayers read the tax law, they form impressions based on their advisors’ reactions to the law. If taxpayers perceive the law as simpler, it may reduce their frustration, and possibly increase compliance. Increased comprehensibility could lead to more efficient tax return preparation and lower cost of compliance. Simplification also might lower government audit cost or more efficiently allocate resources because an understandable law could lead to fewer filing errors and disputes.

See also ERICH KIRCHLER, *THE ECONOMIC PSYCHOLOGY OF TAX BEHAVIOUR* 7 (2007) (“In response to this increasing complexity, many countries have endeavored to simplify the law, although without much success. For instance, New Zealand’s tax law was set into plain English, but still faces the same administrative and compliance problems as before the attempt to make it simpler.”).

what users encounter in the performance of their tax-related tasks, or (2) how successful the outcome is, *i.e.*, by appraising the results assessed by the first type of inquiry.¹⁵⁰ A third form of inquiry, as discussed in Section B, *supra*, concerning how to describe a legal authority, is distinct from the first two forms of inquiry, since it does not directly address taxpayer performance. The first type of inquiry is an empirical assessment, where the challenges are to be exact about the inquiry's objective (to ensure that the assessment provides the information policymakers have need of), and to follow the principles discussed in Section A to ensure the validity of the assessed result. The second type of inquiry, policymakers' appraisal of an assessed result, can rely upon comparisons to judge how good a result is. The first type of inquiry is discussed further in Subsection 1 below, and the second is discussed in Subsection 2.

1. Objectives of an Assessment

Asking what users encounter – what empirical occurrences derive from their performance – would first require a determination of which empirical occurrences, such as resource-spending, frustration or a multifaceted combination of occurrences, to include in the assessment to fulfill the focus of the inquiry.¹⁵¹

A second determination to be made¹⁵² would involve the perspective from which the inquiry will determine whether the performance had a favorable outcome. Will it be a formal perspective, assessing taxpayers' compliance with the existing legal authorities, or will it be the users' perspective, encompassing their own intentions for their performance? Hence, the objective could be either to assess whether users achieved legally correct compliance, or to assess whether they succeeded in reaching whatever goals they had set for their own compliance efforts – which, for some users, might not involve the necessity of getting the task completely right in a legal sense.¹⁵³ From some users' perspective, a less than completely thorough filing might

150. See *infra* III.C.1. and III.C.2.

151. See *supra* Part II.A. (explaining which occurrences the different concepts embrace).

152. See discussion *infra* Part IV.

153. The degree to which taxpayers intend to comply with the tax laws obviously varies, although, as Alex Raskolnikov has expressed it: "Most Americans pay most of their taxes most of the time." Alex Raskolnikov, *Revealing Choices: Using Taxpayer Choice to Target Tax Enforcement*, 109 COLUM. L. REV. 689, 690 (2009).

still represent the successful accomplishment of the performance.¹⁵⁴

A third determination to make in specifying the assessment's objectives would be whether the focus of the inquiry concerns what occurred in an objective sense, or whether it concerns users' perceptions of what occurred. A user's perception might be, for instance, that she or he successfully accomplished her or his goal of filing accurately for a charitable contribution deduction, despite having in fact have taken too high or low a deduction; conversely, the user may believe that she or he failed to accomplish her or his goal of a thorough filing despite having in fact accomplished that very thing. Consequently, the perspective from which the occurrence is viewed – whether the occurrence is factual or perceived – is part of the necessary specification of objectives for the assessment.¹⁵⁵

Hence, the three determinations regarding objectives of the assessment are: first, which kinds of empirical occurrences are considered relevant as indicators of what users encountered, according to policymakers' focus in the inquiry; second, whether the focus of the inquiry pertains to users' lawful compliance or to their ability to meet the goals they have set for themselves, *i.e.*, what they actually are trying to accomplish; third, whether the evaluation will assess what objectively occurred, or users' perception of what occurred.

2. Appraisal of the Assessed Result

Policymakers' appraisal of the results of an assessment of what users encountered while performing their tax-related tasks can be substantiated in various ways, particularly through comparison. Three different comparisons, presented here using compliance cost as an example, could be: (1) "Less is best"; (2) Comparison with standard levels; and (3) Comparison with a relevant alternative.

Less is best. Compliance cost is a "cost," and lower cost is preferable. The "less is best" approach can be understood as a simple look at the magnitude of the cost and appraisal of lower cost as more desirable. This is the premise behind Slemrod's example, and his designation of the terms "simplicity" and "complexity."¹⁵⁶ The "less is best" approach can be applied either

154. See generally *infra* Parts IV.A., and IV.B.1.

155. See generally *infra* Part IV.C. (presenting objective and subjective measures of what happened in the performance).

156. See *supra* Part II.A.1.

from the viewpoint of the individual user segments whose tasks are being evaluated, or on some aggregate, societal level.¹⁵⁷ Under this approach, every change that reduces compliance cost is appraised as an improvement.¹⁵⁸

High or low level. Assessed compliance cost can be converted to a ratio between it and the amount of the revenue or entitlement in question, which can be compared with a standard cost-revenue ratio.¹⁵⁹ Such a comparison could in turn support a rough characterization of the assessed compliance cost as high or low.¹⁶⁰

Relative compliance cost. The designation “relative compliance cost”¹⁶¹ can be applied to refer to a compliance-cost comparison between two alternative scenarios with respect to the matter being assessed.¹⁶² An example is two different legal systems’ schemes for charitable contribution deductions. This kind of comparison allows a legal scheme to be pronounced “most cost-effective” among all those compared. Such comparisons, however, are inevitably intricate.¹⁶³ The challenge involves

157. Taxpayer noncompliance, where a user does nothing or performs only limited tasks, would generally also result in less cost and thus, according to this approach, be considered “best.” This predicament can be resolved by designing the inquiry to appraise the outcome from the formal perspective of fully lawful compliance, rather than from the user’s perspective. As a tax policy objective, compliance cost must generally be understood from the perspective of legally correct compliance.

158. This is not to say that Slemrod adheres exclusively to compliance cost as the sole ground for evaluating tax policy. *See, e.g.,* Joel Slemrod, *My Beautiful Tax Reform*, in TOWARD FUNDAMENTAL TAX REFORM 135 (Kevin A. Hassett & Alan J. Auerbach eds., 2005) (emphasizing that tax policy involves a tradeoff among equity, efficiency, and complexity). Similarly, the SCM approach appraises lower cost with the reservation that other considerations are important as well:

But if the businesses are subjected to expenses through regulation, that could have been avoided, then it is a matter of a societal waste. It is thus important to regulate the businesses’ conduct in a more optimal – and, for the businesses, a less resource intensive – manner, without the overall societal objectives with the regulation being set aside.

SCM MANUAL, *supra* note 30, at 5.

159. *See, e.g.,* Joel Slemrod, *Which Is the Simplest Tax System of Them All?*, in ECONOMIC EFFECTS OF FUNDAMENTAL TAX REFORM 367-68 (Henry J. Aaron & William G. Gale eds., 1996) (arriving at a cost-revenue ratio of around 10 percent, and discussing the difficulties in estimating compliance cost).

160. For a discussion of possible pitfalls and the need for care in interpretations of cost-revenue ratios, *see, e.g.,* CEDRID SANDFORD ET AL., *supra* note 35, at 19.

161. This Article adopts the term “relative” from the terminology of usability to designate comparisons between assessed levels. *See infra* Part IV.D. (explaining origin of the term “relative” in usability terminology, where it concerns the comparison of levels of usability of two products designed to solve the same problem).

162. *See infra* note 251.

163. Comparing tasks instead of legal schemes, however, can make useful comparisons more feasible. Regarding the truthfulness of comparisons of compliance cost, Sandford advanced the following caveat in his latest book. CEDRIC SANDFORD, *WHY TAX*

whether the features of the schemes being compared are “similar” enough for a comparison of the compliance costs originating from each scheme to be fair.¹⁶⁴

3. Possible Elements in an Assessment of Outcome of Use

This Subsection and the following Subsection 4, with the aid of two Figures, provide a graphic representation of how the different elements of comprehensive analyses relate to each other as presented in this Article.¹⁶⁵ The first Figure situates the concepts of compliance cost and psychological cost, whereas the Figure in Subsection 4 depicts the diverse current notions of tax complexity.

In Figure 3 below, the columns represent the elements of an assessment of outcome of use. The long horizontal rectangle and the large arrow at the right illustrate the elements’ connection to the two inquiries addressing taxpayer performance: (1) assessment of what users encounter empirically in their performance of tax-related tasks, and (2) appraisal of the results of that assessment. The arrow at left displays the third form of inquiry, concerning product-oriented descriptions of legal authorities, which is distinct from assessments of empirical occurrences.¹⁶⁶

SYSTEMS DIFFER: A COMPARATIVE STUDY OF THE POLITICAL ECONOMY OF TAXATION 137-38 (2000),

In the light of the international comparisons of administrative costs and of compliance costs illustrated in this chapter, it might seem that such comparisons are likely to mislead rather than reveal the truth and that they are useless. Such a conclusion would be wrong. But it is important that international comparisons of administrative and compliance costs should be regarded as tools to raise questions rather than providing immediate answers. . . . It may well be that, for at least some time to come, policy-makers will get most value from very detailed comparisons of the administrative and compliance costs of small bits of the tax system, such as particular features of a particular tax.

164. The challenges in attaining comparisons that could justify pronouncing one income tax scheme better than another arise from three variables: (1) efficiency and equity effects, (2) variations in compliance obligations among otherwise similar legal schemes, such as the amount of data required for information reporting, and (3) the context in which the performance is undertaken. The smaller the area for true comparisons becomes when the more rigidly “similar” is interpreted. Understanding relative comparisons principally as tools for policymakers might help circumvent the danger of becoming sidetracked by overly scrupulous questions as to whether pronouncing one scheme most cost-efficient is an “unfair” comparison between apples and oranges.

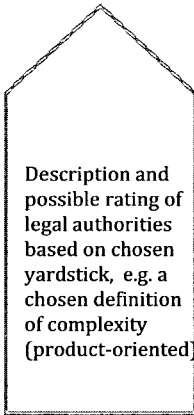
165. See *infra* Figure 3, p. 71 and Figure 4, p. 73. Because the tax complexity field is broader than what is discussed here, the suggested configuration does not in any way purport to be exhaustive.

166. See *supra* introduction III.C. and III.B.

Possible substance for improvements

Assessment (process-oriented)	Basis	Contextual issues	Activities	Outcome of use	Variables	Assessed result
	<ul style="list-style-type: none"> • Legal authorities 	<ul style="list-style-type: none"> • Users • Tasks • Interplay with environment 	<ul style="list-style-type: none"> • Familiarization • Information retrieval • Calculation • Etc. 	<ul style="list-style-type: none"> • Multifarious occurrences: • Time • Agitation • Etc. 	<ul style="list-style-type: none"> • Perspective (formal or users) • Factual (objective) or Perceived (subjective) 	<ul style="list-style-type: none"> • Measured and comprehended as compliance cost, psychological cost, or usability

Assessment setup



Objectives for the assessment

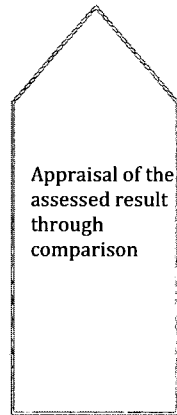


Figure 3¹⁶⁷

The first column, “basis,” illustrates that one or more legal authorities can be either the delineation of a particular assessment, or an object of description and possible rating based on a chosen yardstick. In other words, for a process-oriented assessment, the legal authority, as basis, is an element of the setup, whereas for a product-oriented description, it is the object being described.¹⁶⁸ The last column, “assessed result,” shows that each concept assesses certain aspects of outcome of use.¹⁶⁹ For example, in compliance-cost studies, time and defrayed cost are assessed as economic cost. The assessed results are then the object of appraisals.

4. Tax Complexity Notions with Regard to Assessment of Outcome of Use

Figure 4 below displays which elements of an assessment of outcome of use, or which underlying phenomena, are the focus of the various notions of tax complexity.

167. See *infra* Part III.C.3 and discussion below.

168. See *supra* Parts III.A., and III.B.

169. See *supra* Part II.A., and Figure 2 in Part III.A.

"Complexity estimate": Analysis of attributes of legal authorities through an abstract definition of complexity based on incidence of attributes that the definition regards as complex (product-oriented)

"Complexity experience": Particular occurrences of what happens for users in the performance of their tax affairs, as aspects whose characteristics are uniquely complex in accordance with a defined concept of complexity (process-oriented)

Basis	Context of use	Activities	Outcome of use	Viewpoints	Assessed result
-------	----------------	------------	----------------	------------	-----------------

•Legal authorities

•Users
•Tasks
•Interplay with environment

•Familiarization
•Information retrieval
•Calculation
•Etc.

Complexity denoting linguistic difficulty ("plain English" discussions)	Complexity denoting degree of differentiation in legal authorities (Kaplow)
---	---

Complexity denoting characteristics of various types of activities, e.g., activities for executing compliance obligations or activities for familiarization with written or unwritten rules (Bradford)

Complexity denoting the assessed result if it was appraised as poor (Roberts; possibly Bradford)	Complexity denoting total resource cost (Slemrod; possibly Bradford concerning denoting compliance cost)
--	--

Figure 4

The large arrows relate the notions of tax complexity discussed in Part II.B. to each phenomenon with which they generally are concerned. The bubble at top left, "Complexity estimate," refers to complexity as designating abstract definitions applicable to estimates of *product* attributes.¹⁷⁰ The bubble at top right, "Complexity experience," draws its inspiration from the general complexity literature outside the tax field. The general complexity literature contains additional conceptions of complexity, which designate as "uniquely complex" certain aspects of task performance involving incidents of choice and contingency.¹⁷¹ Such notions of complexity are process-oriented, since they elucidate possible occurrences deriving from performance.¹⁷² The tax complexity literature does not appear to offer any notion of tax complexity pertaining to or defining taxpayer experiences which are "uniquely complex." However, adoption and development in the income tax field of a notion of complexity designating something uniquely complex actually occurring for taxpayers might be a promising approach to designing analyses, which would facilitate a broader and more in-depth understanding of taxpayers' experiences.¹⁷³ This would enable policymakers to ascertain whether the designated complexity occurrence represents a problem to mitigate.

170. See *supra* Part III.B.1.

171. A possible example of a complexity concept designating specific occurrences for individuals is the general idea of hypercomplexity. See QVORTRUP, *supra* note 140, at 6 ("A short definition says that hypercomplexity is complexity inscribed in complexity, e.g., second-order complexity.").

172. See *supra* Part III.B. (explaining the term "process-oriented").

173. Merely relabeling as "complexity" an occurrence, which is usually described using another term, for example, referring to agitation as "complexity," would not contribute to more comprehensive assessments; however, developing a novel conception of complexity, which encompasses something distinct presumably would. Pursuit of an inquiry in the tax context regarding complexity as a unique occurrence for the user is beyond the scope of this Article.

This Part has shown how instructive the usability framework can be, both for obtaining valid assessments by taking "context of use" into account (*i.e.*, basing an assessment of taxpayer experience on a depiction of the real-world context in which the user segment actually performs activities), and also for differentiating the various notions of tax complexity currently found in the literature. Examining these notions together with an overview of the underlying phenomena they embrace explicates a significant reason why "tax complexity" is regarded as elusive: the same term designates numerous dissimilar issues.¹⁷⁴ Moreover, unstated components of an inquiry, such as whether it applies a formal perspective or a user perspective, or whether the assessment concerns what occurred as a factual matter or as a user's perception of what occurred, can only add to the ambiguity. Furthermore, as none of the current tax complexity notions specifically addresses occurrences, which can be understood as "uniquely complex," they offer no special value for such analyses either.

Hence, users' performance of compliance tasks can be analyzed more intelligibly through less ambiguous concepts that offer greater clarity regarding the occurrences they include, their means of measuring, and the way they comprehend assessed results. Among the current concepts, those which best satisfy these criteria are the compliance-cost concept and, as introduced in the next Part, the usability concept.

IV. THE USABILITY CONCEPT

This Part presents the guiding principles for assessments of usability. Whereas Part III utilized the usability *framework* to demonstrate the importance of context of use and to analyze the relationships among the various concepts and notions, this Part introduces the usability *concept* to the income tax field. The usability *concept* both embodies the tax policy objective of enhanced usability, and provides a unique tool for analysis of taxpayers' empirical experience in performing tax-related tasks.¹⁷⁵ Whereas compliance cost, for example, comprehends resource-spending in terms of economic cost, usability comprehends what occurs for users in terms of "quality in use."¹⁷⁶ Hence, this Article puts forth the usability concept as a different

174. See Figure 4 (citing many different examples of what complexity denotes).

175. The compliance-cost concept has a similar duality of functions. See discussion *supra* Part III.C.2.

176. See discussion *infra*. See also *supra* note 9 (explaining "quality in use" as a simplified way to describe what usability embraces).

way to study and ultimately improve the experience of taxpayers in performing tax-compliance tasks.¹⁷⁷ Usability is important to the extent that people performing activities for tax compliance are like product users in general: they value the feeling of smoothness in their performance, dislike failing in the execution of tasks, and judge resource-spending as wasteful if their efforts appear futile or excessive. And the extent of their feelings of ease or anxiety in response to performing their activities is an integral part of how favorably they view the performance. Ben Shneiderman, a pioneer in the field of usability,¹⁷⁸ offers this vivid general description of high usability: “[It] generate[s] positive feelings of success, competence, mastery, and clarity in the user community Creating an environment in which tasks are carried out almost effortlessly and users are ‘in the flow’ requires a great deal of hard work by the designer.”¹⁷⁹ In the income tax context, a desirably high level of usability might be found, for instance, where various user segments, in carrying out activities in order to obtain a charitable-contribution deduction, experience a high degree of mastery, smoothness, and comfort in their respective performance. A usability assessment would embrace multifarious aspects of users’ outcomes of use, such as time and discomfort, which derive from their performance of activities such as familiarization, information retrieval, and calculation.

The usability concept would analyze a broader spectrum of what occurs in users’ performance than any of the concepts currently found in the literature.¹⁸⁰ In fact, it not only would encompass occurrences comparable to those addressed by both compliance cost (resource-spending) and psychological cost (comfort or discomfort), but, in addition, offers a third central aspect of use, “achievement,” which addresses users’ degree of accomplishment of task performance.¹⁸¹ However, because the

177. As a conceptual piece belonging to the realm of discovery, this Article focuses primarily on theoretical development and does not present data for purposes of theory testing. See *supra* note 28.

178. See, e.g., WIKIPEDIA, http://en.wikipedia.org/wiki/Ben_Schneiderman (last visited Oct. 10, 2012) (“He [Ben Shneiderman] also defined the research area of universal usability . . .”).

179. SHNEIDERMAN & PLAISANT, *supra* note 12, at 13. This observation is made with regard to the design of user interfaces, such as those for web pages or computer programs. But there can be little doubt that it would also require a great deal of hard work by a designer of tax systems and administration to enhance usability to the point where users are “in the flow.”

180. See *supra* Part II.A.-B.

181. See *infra* Part IV.B.1.

usability concept would comprehend these occurrences in a different way, it should be seen as complementing the current concepts rather than replacing them. This is illustrated by Figure 3's reference to usability, in the "Assessed result" column, as one of several possible ways to measure and comprehend outcome of use.¹⁸²

The elucidation of broader aspects of what taxpayers encounter in their performance would enable improvement initiatives to be tailored so as to more effectively ameliorate occurrences appraised as reflecting insufficient levels of usability: it would thereby allow higher levels of achievement, less time consumption, more comfort, and the like. Overall, usability as a tax-policy objective would emphasize users' attaining feelings of smoothness and mastery in their performance. This focus might result in a choice of initiatives for improvements different from those chosen where the policy aspiration is merely decreased compliance cost.¹⁸³

A. *Definition of Usability*

ISO defines usability as "the *effectiveness, resource-efficiency, and satisfaction* with which specified users can achieve goals in particular environments."¹⁸⁴ Figure 5 below shows the components of the usability concept.¹⁸⁵ The left side of Figure 5 is identical to Figure 2, and illuminates the components of context of use.¹⁸⁶ The right side of Figure 5, however, now illustrates how outcome of use, i.e., the empirical occurrences deriving from the carrying out of activities for a particular compliance task, is assessed with the usability concept, through the three measures of usability: effectiveness, resource-efficiency, and satisfaction.

182. See *supra* Figure 3.

183. See *infra* Part IV.D.2. (discussing focus of usability improvements).

184. See ISO 9241-11, *supra* note 12, at § 3.1. See also Hornbæk, *supra* note 9, at 79, 82. ISO uses the term "efficiency," but this Article employs the phrase "resource-efficiency" to avoid confusion with the common tax-policy term "efficiency." See *id.*

185. See *infra* Figure 5.

186. See *supra* Figure 2 and Part III.A.

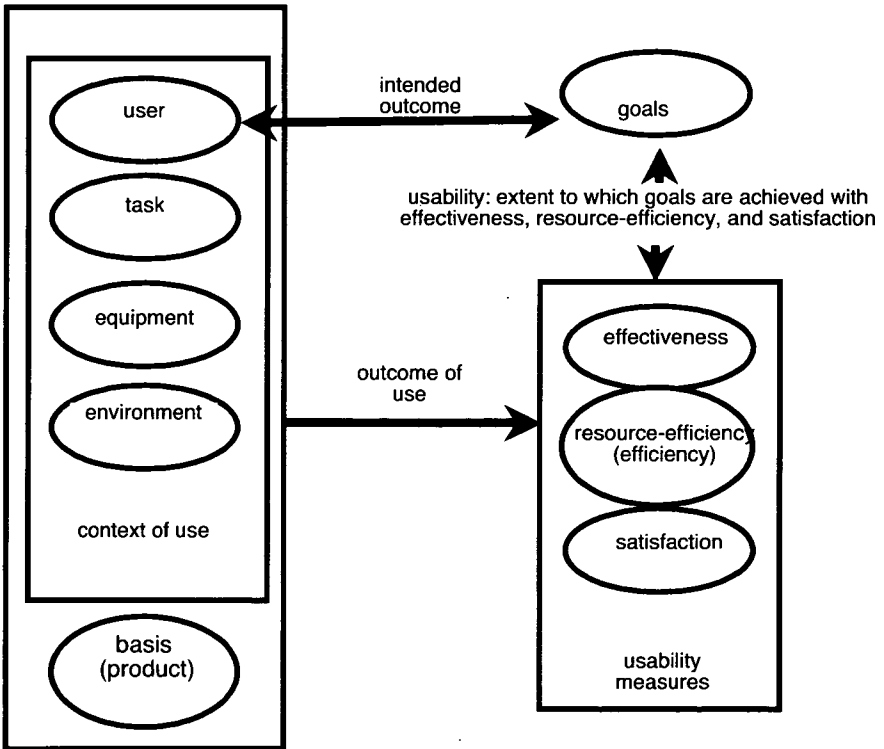


Figure 5¹⁸⁷

An example of a usability assessment in the income tax context could focus on users, who want to take a fully lawful deduction for a charitable contribution. The assessment as to whether these users can correctly accomplish the various activities necessary to obtain such a lawful deduction constitutes the *effectiveness* measure.¹⁸⁸ How many resources, in the form of time, effort, and money, the users have to spend to reach whatever measured degree¹⁸⁹ of effectiveness they achieve would be measured as *resource-efficiency*.¹⁹⁰ Finally, whether the process of carrying out the task was positive or negative – as determined by the user's response to the use, in the form of comfort or stress – would be assessed by the *satisfaction* measure.¹⁹¹

ISO provides the following examples of how the three measures of usability can be operationalized:

Usability objective	Effectiveness measures	Resource-efficiency measures	Satisfaction measures
Overall usability	Percentage of goals achieved	Time to complete a task	Rating scale for satisfaction

187. See ISO 9241-11, *supra* note 12, at § 5.1.2. Figure 5 is taken from ISO, and original ISO designations that have been changed appear in parentheses. *Id.*

188. See further discussion *infra* Part IV.B.1.

189. ISO 9241-11, *supra* note 12, at §§ 3.10& 3.11 (describing "measure" as "Measure (noun): Value resulting from measurement and the process used to obtain that value.").

190. See further discussion *infra* Part IV.B.2.

191. See further discussion *infra* Part IV.B.3.

	Percentage of users successfully completing task	Monetary cost of performing the task	n [e.g., a scale from 1 to 5 which the user applies to rate satisfaction]
--	--	--------------------------------------	--

Table 1¹⁹²

192. ISO 9241-11, *supra* note 12, at Table B.1. Examples of measures less applicable in the income tax context are omitted. In the usability concept's original domain of information and communication technology, the three measures could be illustrated by an assessment of the use of LexisNexis to find articles. The task for evaluation could require a particular segment of users to find ten specific articles through LexisNexis database. Operationalizing the effectiveness measure by assessing what percentage of the goal was achieved might show that, on average, users find eight out of ten articles. Another means of operationalizing effectiveness, either in addition to or instead of the previous operationalization, could assess the percentage of users successfully completing the task; this assessment might show that 70% of users find all ten articles. Operationalizing the resource-efficiency measure, by assessing time, could show that the average time spent on the task before completing it or giving up was one hour. Operationalizing the satisfaction measure, by applying a rating scale, might show that the average user response was 3.5 on a 5-point scale, indicating that users were tolerably comfortable working with the LexisNexis product. Of course, this example is simplified in comparison to a real evaluation. Overall usability would be a qualitative mixture of these results.

The level of usability can be determined either against absolute criteria or against a reference system. See ISO 9241-1, *supra* note 102, at § 4.2. Nielsen offers this explanation:

To determine a system's overall usability on the basis of a set of usability measures, one normally takes the mean value of each of the attributes that have been measured and checks whether these means are better than some previously specified minimum. Since users are known to be very different, it is probably better to consider the entire distribution of usability measures and not just the mean value. For example, a criterion for subjective satisfaction might be that the mean value should be at least 4 on a 1-5 scale; that at least 50% of the users should have given the system the top rating, 5; and that no more than 5% of the users gave the system the bottom rating, 1.

In the legal context, when an assessment of outcome of use includes the aspect of achievement, the assessment must specify the perspective from which achievement is viewed: achievement from the users' perspective, or achievement from a system perspective concerned with full compliance with the law.¹⁹³ A difference will arise between these two perspectives if, for example, the user segment in question does not have full lawful compliance as its goal. A user segment might consist of people who are attempting to accomplish "more or less lawful" compliance, rather than fully lawful compliance. Variations in intended outcomes among users are encompassed by the usability concept's goal component, as Figure 5 illustrates.¹⁹⁴

B. Usability Measures

The ISO 9241-11 usability concept is process-oriented.¹⁹⁵ It is not about describing a product or legal authority; rather, it is about assessing the occurrences deriving from either the *use* of the product or the *performance* of the activities involved in executing tasks induced by the legal authority – in either case, within the relevant context of use.¹⁹⁶ Hence, the starting point is that users carry out activities for the execution of tax-related tasks, and as a result of these activities, empirical phenomena occur, such as time consumption.¹⁹⁷ The three usability measures – effectiveness, resource-efficiency, and satisfaction – are tools for assessing such occurrences; each of them measures particular occurrences, and together the resulting measurements

NIELSEN, USABILITY ENGINEERING, *supra* note 29, at 27.

193. See *supra* Part III.C.1.

194. See *supra* Figure 5. See also ISO 9241-11, *supra* note 12, at §§ 3.8, 5.2.

195. See, e.g., Abran et al., *supra* note 20, at 326.

196. See ISO 9241-11, *supra* note 12, at Annex D, D.1. See also Abran et al., *supra* note 20, at 331 ("The measurement of usability is a complex interaction between users and context of use; this might produce different levels of usability performance for the same product when it is used in the different contexts."). ISO notes that the term "usability" is sometimes encountered as referring more narrowly to the attributes of a product that make it easier to use: "a set of attributes . . . which bear on the effort needed for use and on the individual assessment of such use by a stated or implied set of users." See ISO 9241-11, *supra* note 12, at Annex D, D.1. Such an application of the term "usability," however, pertains to a product-oriented standard, as the description of the attribute of the product is independent of context of use. ISO 9241-11 objects to such application. *Id.* See also Abran et al., *supra* note 20, at 326 (concerning the contrast between process- and product-oriented usability standards).

197. Time consumption is thus an example of the observable fact that when people carry out something, then something happens as a result - at a minimum, an expenditure of time.

can be comprehended as a level of usability.¹⁹⁸ The following Subsections discuss application of the three usability measures.¹⁹⁹

1. The Effectiveness Measure

The effectiveness measure, simply put, assesses whether users can accomplish the attempted task.²⁰⁰ As such, it can be seen as the most basic measurement of how well or poorly a task went, that is, the extent to which users succeeded in accomplishing what they attempted.²⁰¹ ISO defines “effectiveness” as the “accuracy and completeness with which users achieve specified goals.”²⁰² Accuracy and completeness are thus ways to operationalize the effectiveness measure.²⁰³ By assessing accuracy and completeness, the effectiveness measure determines how close taxpayers came to their goals or intended outcomes, thereby elucidating the occurrence of “achievement”.²⁰⁴

198. See *supra* note 192 and accompanying text.

199. Quantified measurements made through these three measures – effectiveness, efficiency, and satisfaction – are an integral part of the definition of usability. Hornbæk puts it very explicitly:

[W]hat we mean by the term usability is to a large extent determined by how we measure it Thus, measures of usability serve to make the general and somewhat vague term usability concrete and manageable. . . . Through operationalization of the usability construct, we find aspects of usability that can be measured.

Hornbæk, *supra* note 9, at 79-80. This corresponds to the situation for compliance cost. For example, what is accounted for as “compliance cost” is determined by what is measurable in terms of “economic cost”. See discussion *supra* Part II.A.

200. See JORDAN, *infra* note 201, at 18.

201. See PATRICK W. JORDAN, AN INTRODUCTION TO USABILITY 18 (2002). (“The most basic measure of whether or not a product is effective for a particular task is whether or not the user can complete that task with the product.”).

Usability concerns use or performance – not the full “consumer value/cost.” For instance, if the product, which is the basis for a usability evaluation, is a lottery ticket, what is being assessed is not whether the user had success with her or his purchase by winning a prize, but simply the use of the lotto coupon. That includes, among other aspects, whether the user accomplished the filling out of the lotto coupon according to her or his goal. That is in exactly those spaces, which the user intended to fill it out. But, the query of whether the user’s selected numbers were the lucky numbers would not be a usability question. See generally ISO 9241-11, *supra* note 12, at Introduction, page iv.

202. ISO 9241-11, *supra* note 12, at § 3.2.

203. See *supra* note 199 (explaining operationalization of “usability”).

204. Conceptually, accuracy can be understood as a measure of whether the user achieved the quality attempted, while completeness can be understood as a measure of whether she or he achieved the attempted quantity. ISO describes them as follows:

To measure accuracy and completeness it is necessary to produce an operational specification to the specified criteria for successful goal achievement. This can be expressed in terms of the quality and quantity of output Accuracy can be measured by the extent to which the quality of the output corresponds to the

For example, for a segment of taxpayers whose intended outcome is to claim no less and no more than a full deduction for their charitable contributions, the achievement aspect would concern whether they accomplished this goal. An effectiveness measure operationalized as accuracy could address whether the computation of the deduction was correct, and an effectiveness measure operationalized as completeness could address whether a taxpayer succeeded in claiming the deduction for the totality of her or his deductible contributions.

As noted, the effectiveness measure of level of achievement makes its assessment according to the user segment's specified goals, or intended outcomes.²⁰⁵ This is relevant to an elucidation of users' abilities, since only by measuring against what users are actually trying to accomplish, i.e., having a user perspective, can an assessment illuminate users' abilities to succeed in their performance. For example, one user segment might not be troubled about whether their performances allow them to receive slightly less or slightly more than the maximum deduction to which they are entitled for charitable contributions. For such a user segment, the effectiveness measure indicates the degree of accuracy and completeness by which their intended outcome is accomplished. Thus, if such a user reached her or his intended outcome of anything from slightly less to slightly more than the maximum deduction, this would be assessed as the user's goal having been achieved. Measuring according to intended outcome thus ensures a more apt elucidation of users' ability to accomplish their goal. It is important to note, however, that as a tax policy objective, usability should generally be understood to embrace user segments whose intended outcome is fully lawful compliance.²⁰⁶

specified criteria, and completeness can be measured as the proportion of the target quantity, which has been achieved.

ISO 9241-11, *supra* note 12, at Annex B, B.5.1.

205. See ISO 9241-11, *supra* note 12, at § 3.2. (defining the term effectiveness).

206. Fully lawful compliance as the point of reference, when usability is the tax policy objective, corresponds to common practice for the compliance-cost concept. See *supra* note 157.

In a usability assessment, variations in degree of intended compliance are incorporated by segmenting users according to goals. A compliance-cost assessment also has to take a position concerning whether it should be based on fully lawful compliance or actual compliance. See, e.g., SCM MANUAL, *supra* note 30, at 17 (explaining "To measure businesses' administrative costs by following a set of regulations, it is important to make clear assumptions about compliance. The costs of full or partial compliance may be measured.").

2. The Resource-Efficiency Measure

The resource-efficiency measure assesses resources spent, and does so in relation to effectiveness.²⁰⁷ The resource-efficiency measure is based on the supposition that, all else being equal, an expenditure of fewer resources in relation to a given degree of effectiveness measurement constitutes a better performance.²⁰⁸ ISO defines resource-efficiency as the resources expended in relation to the “accuracy and completeness with which users achieve goals.”²⁰⁹ The level of achievement obtained by virtue of the resource-spending is thus an integral part of the resource-efficiency measure. For instance, two hours spent performing the activities necessary to claim a deduction for charitable contributions will have been better spent if the user has fully accomplished the intended task than if she or he was able to do so only partially.

The kinds of resource-spending assessed through the resource-efficiency measure can be diverse. ISO names the following: “Relevant resources can include mental or physical effort, time, materials or financial cost. For example, human efficiency could be measured as effectiveness divided by human effort, temporal efficiency as effectiveness divided by time, or economic efficiency as effectiveness divided by cost.”²¹⁰ ISO also refers to the possible measurement of workload, including expenditure of both physical and cognitive resources.²¹¹

The focus of the inquiry should determine the number of aspects of resource-spending that are to be measured. For instance, if the purpose of an inquiry is to assess the overall level of usability in carrying out the activities involved in a charitable-contribution deduction for taxpayers who do not employ tax

207. ISO 9241-11, *supra* note 12, at § 5.4.3.

208. See Hornbæk, *supra* note 9, at 87 (“In the ISO definition of usability . . . time is considered a resource of which successful interfaces minimize consumption. However, in a handful of studies [e.g. for computer games] higher completion times are considered as indicators of motivation, reflection, and engagement.”) This Article believes the welfare implications of compliance tasks to be best understood in line with the main assumption of the ISO definition – and definitely differently from computer gaming.

For a discussion of compliance tasks as an opportunity for taxpayers to consider fiscal responsibilities, see Joseph J. Thorndike, stating, “It’s important to keep people tuned in to taxes. But we can accomplish that goal while still easing the process of completing a tax return. ReadyReturn does exactly that.” Joseph J. Thorndike, *Why Everyone Should Like ReadyReturn—Even the Tax Foundation*, TAX.COM (Oct. 8, 2009, 12:18 PM), <http://www.tax.com/taxcom/taxblog.nsf/Permalink/JTHE-7WMJ94?OpenDocument>.

209. ISO 9241-11, *supra* note 12, at § 3.2.

210. *Id.* at § 5.4.3.

211. *Id.* at Annex B, B 5.2.1.

preparers, time and financial cost would be the obvious clear choice of aspects to measure. Adding more measurements, such as a human effort measurement based on users' energy expenditure in completing the workload, would yield a more comprehensive assessment of the outcome of use.

3. The Satisfaction Measure

The satisfaction measure assesses users' comfort and attitudes, i.e. broadly speaking, the users' psychological well-being, in response to performing the activities necessary to complete tasks.²¹² If, for instance, performing the activities causes exasperation, this will be assessed by the satisfaction measure.²¹³ ISO describes the satisfaction measure as follows:

Satisfaction (defined as freedom from discomfort, and positive attitudes towards use of the product) is a response of users to interaction with the product. . . . Objective measures can be based on observation of the behavior of the user (e.g. body posture, body movement, frequency of absences) or can be based on monitoring the physiological responses of the user.²¹⁴

Hence, the satisfaction measure would assess taxpayers' responses to managing the tasks induced by the legal authority that is the basis for the assessment. Such assessment of the psychological occurrences deriving from taxpayers' performance of a compliance task should not be conflated with an inquiry into the legal authority's popularity or perceived legitimacy with various segments of taxpayers, which would involve, for instance, questions of whether or not taxpayers sympathize with the inclusion of charitable-contribution deductions in the tax system.²¹⁵ Questions about support or opposition to specific tax

212. See discussion *infra* Part IV.B.3.

213. See discussion *infra* Part IV.B.3.

214. ISO 9241-11, *supra* note 12, at Annex B, B.6.

215. *The IRS Oversight Board's 2010 Taxpayer Attitude Survey* is an example of the kind of survey that, as to the majority of its questions, should not be conflated with an assessment of taxpayers' responses to performing tax-related activities. See *The IRS Oversight Board's 2010 Taxpayer Attitude Survey*, IRS OVERSIGHT BOARD, <http://www.treasury.gov/irsob/reports/2011/IRSOB%202010%20Taxpayer%20Attitude%20Survey.pdf>. For instance, when taxpayers respond to a question like "It is every American's civic duty to pay their fair share of taxes," the answers do not provide information about taxpayers' responses to performing tasks. (The 2010 IRS Oversight Board reported that 69% completely agree.) See *id.* at 3. However, one of the survey's questions can, to a certain extent, be understood as referencing taxpayers' responses to use. The question is as follows: "Most people have had some type of interaction with the

provisions, or to the income tax in general, are not part of an analysis of how well or how poorly taxpayers perform compliance tasks, as they do not specifically relate to performance of an induced task.²¹⁶ That having been said, attitudes toward the tasks might influence users' responses to performing their activities, such as whether they become agitated or remain calm, and thus affect the occurrences assessed by the satisfaction measure.²¹⁷ For instance, taxpayers carrying out activities for the charitable-contribution deduction might be less inclined to experience stress because they feel good about having made the contribution and getting a deduction, whereas taxpayers executing a task to which they feel strongly opposed might be more likely to experience agitation. Hence, the satisfaction measure can indirectly embrace factors such as sympathy or resistance when they affect users' responses to their performance.

The satisfaction measure is commonly operationalized through the application by users of a rating scale (e.g., a scale from 1 to 5).²¹⁸ By this or other means for operationalization, the satisfaction measure can really offer an assessment of psychological occurrences by showing an estimate of lower or higher levels of comfort and the like. Hence, in contrast to the prevalent approach to the concept of psychological cost, the satisfaction measure does not attempt to transform psychological well-being into monetary terms.²¹⁹ Quite different, the satisfaction measure proposes another way to measure "the

IRS, whether it's just filing your tax return or actually speaking with an IRS representative. How satisfied would you say you have been with your personal interaction with the IRS?" *Id.* at 14 (reporting 38% very satisfied; 39% somewhat satisfied; 9% not very satisfied; 6% not at all satisfied; and 9% not knowing, having no answer, or not responding).

216. See, for example, Nielsen, who makes a comparable point in regard to computers:

Note that the notion of subjective satisfaction as an attribute of usability is different from the issue of the public's general attitudes towards computers. Even though it is likely that a person's feelings toward computers as a general phenomenon will impact the extent to which that person likes interacting with a particular system, people's attitudes toward computers in general should probably be seen as a component of the social acceptability of computers rather than their usability.

Nielsen, *supra* note 29, at 33.

217. See generally *id.*

218. See *id.* at 34; See also discussion *infra* Part.IV.C. (explaining this as a subjective measure, and designating a rating scale also as an "attitude scale").

219. See *supra* note 58 and accompanying text (noting that the concept of psychological cost has the goal of estimating psychological well-being in monetary terms, but has not been able to realize it).

anxiety, frustration, and anger associated with filing", to quote Joseph Bankman regarding what the compliance cost and psychological cost concepts have not managed to measure as "cost".²²⁰

The satisfaction measure is assessed directly, in its own right, and is not inferred or otherwise generated from other occurrences.²²¹ In this way, it differs from the assessment of psychic cost referred to in the discussion of the psychological cost concept.²²² In the study Sandford cites, elderly taxpayers' perception was that they had not been able to fulfill their goal of a fully lawful and thorough filing.²²³ Mainly based on this aspect of their performance, the study inferred that these taxpayers incurred high psychological cost.²²⁴ However, although perception of a low level of achievement will likely elicit discomfort for many people; as an analytical matter, there is no necessary relationship between level of achievement and psychological well-being, which are two distinct occurrences.²²⁵ Perceived level of achievement is not uniformly correlated to the occurrences measured by satisfaction, as users' psychological responses to their perceptions of achievement might vary significantly.²²⁶ For instance, some users might have a response of high agitation, while others might hardly be affected. The satisfaction measure handles this issue by assessing the user's response directly and independently, instead of inferring it from taxpayers' other encounters.²²⁷

For some tasks, users' responses to the performance, as measured by satisfaction, might account for a significant share of what matters in the overall level of usability.²²⁸ If, for instance, it were demonstrated for a specific task that levels of achievement and resource-spending, as measured by effectiveness and resource-efficiency, were reasonable, but that level of user

220. See *supra* note 52.

221. See the description of the satisfaction measure in this Subsection and the example of its operationalization in Table 1. See *supra* Part IV.A.

222. See *supra* Part II.A.2. and note 55.

223. Whether these users actually had or had not accomplished fully lawful compliance is apparently not known. See discussion *infra* Part VI.C. (discussing the application of objective and subjective measures).

224. SANDFORD ET AL., *supra* note 35, at 18 ("Psychic costs are particularly felt by the old and retired . . .").

225. See Hornbæk, *supra* note 9, at 94 (discussing correlations between measures).

226. See *id.*

227. See the description of the satisfaction measure in this Subsection and the example of its operationalization in Table 1. See *supra* Part IV.A.

228. See generally ISO 9241-11, *supra* note 12, at Annex B, B.3 (emphasizing that "care should be taken that appropriate weight is given to each measurement item.").

comfort was low, improvement initiatives could be focused on steps to enhance users' psychological well-being.

C. *Objective and Subjective Measures*

What occurs for users in their performance of tasks can be measured either objectively or subjectively.²²⁹ This choice of viewpoint for the measures relates to whether the focus of the inquiry pertains to what happened factually or to what happened in the users' perception.²³⁰ Objective measures do not depend on user perception: the researcher can obtain, discuss, and validate these measures²³¹ in ways not possible with subjective measures.²³²

ISO describes subjective measures in regard to the satisfaction measure as follows:

Subjective measures of satisfaction are produced by quantifying the strength of a user's subjectively expressed reactions, attitudes, or opinions. This process of quantification can be done in a number of ways, for example, by asking the user to give a number corresponding to the strength of their feeling at any particular moment, or by asking users to rank products in order of preference, or by using an attitude scale based on a questionnaire [e.g., ranking on a scale of 1 to 5].²³³

A subjective measure in regard to the satisfaction measure for assessment of the comfort or discomfort occurring for users in response to performing the activities necessary to claim a charitable-contribution deduction might use an attitude scale, thereby illuminating the user segment's response as the users themselves understood it.²³⁴ By contrast, an objective measure in regard to the satisfaction measure might involve observing users' behavior or monitoring their physiological responses.²³⁵ However, users' psychological response to the performance is not the only occurrence that can be measured both objectively and

229. See Hornbæk, *supra* note 9, at 91. Hornbæk points out that the differentiation between objective and subjective measures is not a substantial epistemological distinction. *Id.*

230. See *id.* at 91.

231. See *supra* note 189 (explaining the ISO term "measure").

232. See Hornbæk, *supra* note 9, at 91.

233. ISO 9241-11, *supra* note 12, at Annex B, B.6.

234. See *id.* at B.6.

235. *Id.*

subjectively.²³⁶ A subjective measure for resource-efficiency measure concerning time spent would assess duration as perceived by users instead of the objective length of time elapsed.²³⁷ Taxpayers' perception of time expenditure could be lower than, similar to, or higher than what the objective measurement would show.²³⁸

A subjective measure for effectiveness measure, assessing perceived level of achievement, is possible as well.²³⁹ This perception can be measured by asking users questions relating to their confidence in the accuracy and completeness of performance.²⁴⁰ Here, user perception could show variation in two directions, with some users believing they had accomplished a higher level of achievement than the one measured objectively, and others believing they had accomplished a lower level.²⁴¹ Some users' perceptions, of course, would correspond fairly well to the objective measurement.

236. See, e.g., *id.* at Annex B, B.4("[S]atisfaction can also be inferred from objective measures of the behavior of the users, and estimate of effectiveness and efficiency can also be derived from the subjective opinions which the users express about their work and its outputs.").

237. See Hornbæk, *supra* note 9, at 92.

238. Another example is that resource-spending due to time spent waiting on the phone before reaching the competent person on a revenue body's service line can be measured either as the objective length of time spent or as the experienced duration. See Jacob Hornik, *Subjective vs. Objective Time Measures: A Note on the Perception of Time in Consumer Behavior*, 11 J. CONSUMER RES. 615, 617 (1984). Waiting time is an area in which it is well documented that experienced duration is significantly longer than objective duration. See, e.g., *id.* ("Results show that individuals exhibit a tendency to overestimate waiting time."); Rabi G. Mishalani et al., *Passenger Wait Time Perceptions at Bus Stops: Empirical Results and Impact on Evaluating Real-Time Bus Arrival Information*, 9 J. PUB. TRANSP. 89, 103 (2006) ("[P]assengers perceive waiting times to be greater than actual waiting times . . .").

239. See, e.g., ISO 9241-11, *supra* note 12, at Annex B, B.4.

240. See, e.g., *id.*

241. The salience of discrepancies between what taxpayer perceive about their handling of tax-related tasks and an objective view of how they managed them might be greater for the income tax, given its character as a self-reporting system, than it would be in the case of legal authorities that induce users to make an application and afterwards receive an administrative ruling. Audit rates in general are very low, so auditing will only rarely rectify such disparities, and tax bodies generally do not confirm that an induced task is accomplished fully lawfully. See John Tozzi, *The Taxman Cometh: Keep Your Head (and Revenues) Down*, BLOOMBERG BUSINESSWEEK, http://images.businessweek.com/mz/11/04/1104_46sbtaxes.pdf (last visited October 10, 2012) (hereinafter "*The Taxman Cometh*"). Therefore, for instance, users who intended to comply lawfully but doubt whether they accomplished this goal must live with this uncertainty, as they do not receive any affirmation of how well they succeeded as a factual matter – unless they actually did not achieve lawful compliance *and* were among the few who were audited. See *supra* Part IV.A.

The average likelihood of a U.S. individual being audited is around 1 percent. See *The Taxman Cometh* (referencing IRS data).

Kasper Hornbæk has proposed a working model of usability measures that encompasses both objective and subjective measures.²⁴² The Figure is based on the ISO standard²⁴³ and is helpful for further illustration of the usability measures.

Usability aspects	Objective measures	Subjective measures
-------------------	-----------------------	------------------------

242. See Hornbæk, *supra* note 9, at 96.

243. See *id.*

Outcomes (effectiveness)	Expert assessment, comprehension	Users' perception of outcome
Interaction process (efficiency)	Time, usage patterns, learnability	Subjectively experienced duration, mental workload, perception of task
Users' attitudes and experiences (satisfaction)	Physiological usability, reflex responses	difficulty Validated questionnaires

Figure 6²⁴⁴

When users perform activities for tax-related tasks, both the objective and subjective measures of what occurred are valid, but they represent different kinds of insights.²⁴⁵ If the focus of the inquiry concerns how users perceive their compliance efforts, subjective measures are advantageous,²⁴⁶ whereas if the focus

244. See Hornbæk, *supra* note 9, at 96. Figure 6 is a simplified version of Hornbæk's figure, omitting illustrations of some of Hornbæk's discussions and suggestions for improvement of the ISO usability concept. See *id.*

245. See *id.* at 92 (explaining that evaluations with either objective or subjective measures may lead to different conclusions regarding usability).

246. An example of an inquiry where subjective measures might be advantageous is the Danish government's so-called "Burden-Hunters" project. MindLab, THE BURDEN-

concerns how well users perform in fact, then objective measures are appropriate.²⁴⁷

D. Usability Improvements

The following subsections describe approaches to appraisal of the result of a usability assessment, and potential areas of focus for improvement when usability is the objective.

1. Appraisal of Assessed Level of Usability

Policymakers' appraisal of a measured level of usability can be substantiated through different comparisons. Part III.C.2. discusses these methods of substantiation: (1) "Less is best," (2) Comparison with standard levels, and (3) Comparison with a relevant alternative. These comparisons are also informative when the question is how to appraise the result of a usability assessment.²⁴⁸

"Less is best," as an approach to appraisal, expresses the basic notion that from the individual taxpayer's viewpoint, users are better off when they are free of the need to perform a task.²⁴⁹ Consequently, when thinking about usability generally as an aspiration for users' welfare, the "less is best" approach would consider removal of a task an improvement.²⁵⁰

HUNTER TECHNIQUE, A USER-CENTRIC APPROACH TO CUTTING RED TAPE, 4 (2008), http://mind-lab.dk/assets/344/The_burden-hunting_technique.pdf (last visited Oct. 10, 2012). The purpose of this project is to go further than the Standard Cost Model assessment of compliance cost through increasing the emphasis on the burden experienced by enterprises – the burden of irritation – in other words, an enterprise's subjective experience of satisfying the demands imposed by official business regulations. . . . The burden hunters supplement existing red-tape reduction efforts by placing particular emphasis on the burdens experienced by enterprises, and on how other factors besides the expenditure of time can cause enterprises to regard business as being a burden. *Id.* Distinguishing between objective and subjective experiences of time is a possible way to assess how well users see themselves as performing. In general, applying subjective measures appears to be a potential method to capture aspects of the outcome of use that would be adequate indicators for the project's research questions.

247. See, e.g., Hornbæk, *supra* note 9, at 92.

248. See *supra* note 192 (explaining how to determine level of usability).

In the domain of information and communication technology, common methods of appraising level of usability are either comparison against absolute criteria or comparison against a reference system. See ISO 9241-1, *supra* note 102, at § 4.2. Obviously, such standard levels for comparison are not yet available in the income tax field.

249. See *supra* Part IV.B.2. and accompanying note 208.

250. With regard to assessment of performance, there is, strictly speaking, no usability issue when there is no performance to assess. Still, the "less is best" means of comprehension can be justified if the assessment that is envisaged includes several tasks or considers the omission of a previously executed activity. In a case where one task or

The standard way to appraise usability is to make a comparison with a relevant alternative to determine which is "most usable."²⁵¹ An example is a comparison of performances from two different legal systems' schemes for charitable-contribution deductions.²⁵² This approach to appraisal is also designated as "relative usability."²⁵³ The challenges in the tax context for relative usability comparisons are the same as those for relative compliance cost.²⁵⁴

An example concerning the deduction for charitable contributions in Denmark is illustrative. Before 2008, taxpayers were required to file for charitable-contribution deductions.²⁵⁵ Since then, the charitable associations entitled to receive tax-deductible donations have had to submit information to the revenue body about the amounts received from each taxpayer, as a requirement for the taxpayer to enjoy the right to the deduction.²⁵⁶ The deductible amount then appears on the pre-filled tax returns that Danish taxpayers receive from the revenue body.²⁵⁷ The taxpayer's only task is to inform the charitable

activity (as to which users often did not reach a high level of achievement) were eliminated, this would raise the assessed overall level of usability.

See also Slemrod, *My Beautiful Tax Reform*, *supra* note 169, at 13 (regarding the additional concern in tax policy for equity and efficiency).

251. See, e.g., Jacob Nielsen & Victoria L. Phillips, *Estimating the Relative Usability of Two Interfaces: Heuristic, Formal and Empirical Methods Compared*, 24-29 INTERCHI 93 214, 214 (1993).

As always when employing the ISO usability concept, which is process-oriented, it is the specific context-dependent *uses* of a product, or *performances* of a task induced by the basis of the assessment, that are assessed, so it is the levels of usability of the respective uses or performances which are being compared. See e.g. Abran et al., *supra* note 20, at 331; see also *supra* Part IV.B. and accompanying note 196.

252. What is true for compliance cost also holds for usability: that the more flexibility there is regarding what are deemed to be acceptably similar tax schemes, the more room there is for comparisons and for determinations of "most usable" among the different alternatives. See *supra* note 164. Acceptance of the "similarity" of two or more tax schemes will thus confirm that it makes sense to compare those alternatives' levels of usability.

253. See, e.g., Nielsen & Phillips, *supra* note 251, at 214 ("One often needs to assess the relative usability of two or more user interfaces, designed to solve the same problem.").

254. The existence of three variables (efficiency and equity effects, variations in compliance obligation, and context of use) makes formation of precise judgments from the comparisons difficult. See *supra* note 164. Yet, as a tool for policymakers, relative usability comparisons would be valuable.

255. SKAT, THE DANISH TAX AND CUSTOMS ADMINISTRATION, Announcement of January 11, 2008 (describing the change).

256. Ligningsloven (LBK nr. 1017 of October 28 2011) (Danish Tax Assessment ACT) § 8 A.

257. See SKAT, THE DANISH TAX AND CUSTOMS ADMINISTRATION, CHARITABLE ASSOCIATIONS' DECLARATION OF DEDUCTIBLE GIFTS AND DONATIONS, Dec 20, 2011, <http://www.skat.dk/SKAT.aspx?oId=1797587&vId=204330> (last visited Oct.10, 2012).

association of her or his name and personal identification number, a number Danes know by heart.²⁵⁸ Presumably, taxpayers enjoy a high level of usability in performing this task, as it would appear that not much could go awry in executing it. Before the change, however, the levels of achievement were lower and resource-spending higher.²⁵⁹ Likewise, under the new rules, the charitable associations might also have a high level of usability in performing their task of submitting information, as they would generally have the expertise and data systems to master and execute the necessary activities smoothly.²⁶⁰ Thus, it is reasonable to judge the change as a usability improvement over the old alternative, assuming that the effectiveness measures (*i.e.*, accuracy and completeness with which users achieve specified goals) show a high level of achievement, that the resource-efficiency measure shows that the amount of resources expended in relation to accuracy and completeness is low, and that the satisfaction measure shows a high level of comfort in carrying out the activities for the task. A high level of comfort in response to the performance is to be expected when users have been able to perform well, as is presumed in the example here.²⁶¹

2. Focus of Usability Improvements

When usability is the objective of improvements, the focus falls predominantly on the occurrences indicating users' accomplished level of achievement, on the quality of users' experience, and on how users perceive their performance (as assessed through subjective measures).²⁶² High levels indicate high usability. The example above, of a new scheme for charitable associations to submit information regarding donors' contributions, would likely fulfill all of these objectives for

258. *See id.*

259. *See* SKAT, THE DANISH TAX AND CUSTOMS ADMINISTRATION, COMPLIANCE WITH THE TAX RULES BY PRIVATE INDIVIDUALS IN DENMARK: TAX YEAR 2006 3, 5-6 (2009) [hereinafter "COMPLIANCE WITH THE TAX RULES BY PRIVATE INDIVIDUALS IN DENMARK"]. In tax year 2006, before the new rules came into force, 7.9 percent of the taxpayers who deducted contributions to charitable associations had a disparity between what they reported and fully lawful compliance. *Id.* at 49. The extent to which the disparity was intentional or unintentional is uncertain. *See id.* at 10-11. Under the new rules, disparities caused by taxpayers' task management are much less likely.

260. The deduction for charitable contributions in Denmark is reserved for contributions to approved associations, *see supra* note 256, at § 8.A.

261. *See supra* Part IV.A.3. (discussing the relationship between level of achievement and psychological well-being).

262. *See, e.g.*, SHNEIDERMAN & PLAISANT, *supra* note 12, at 13 (describing high levels of usability as enabling users to be "in the flow").

enhanced usability. By contrast, had the example been appraised with nothing other than compliance cost as the objective, then only lower resource-spending, by the donors and the charitable associations combined, would have allowed the initiative to be deemed an improvement.²⁶³

The work toward improvements might be constrained by politically-entwined requirements about maintaining current tax schemes.²⁶⁴ However, even with restrictions requiring that tax schemes remain similar in scope, usability can be improved, primarily through changes in context of use, as in the Danish scheme for deductions for charitable contributions. In principal, each component of context of use can be considered for improvement initiatives.²⁶⁵

The most promising source of changes in context of use that could lead to continued improvements in outcomes of use appears to be the increasing application of information technology.²⁶⁶ According to OECD, the main types of electronic services now offered by most revenue bodies to both taxpayers and tax professionals include:

- Access to a comprehensive range of tax and other information via the Internet;
- Electronic filing of tax returns;
- Fully and/or partially pre-filled tax returns;
- A mix of electronic payment facilities for all taxes;
- Access to personal taxpayer information via an online taxpayer “account”; and
- Modern telephone call centers to provide more accessible services.²⁶⁷

263. See generally *supra* Part III.C.2.

264. See, e.g., Daniel N. Shaviro, Principles for Comprehensive Income Tax Reform, Testimony Before the United States Senate Committee on Finance, 1, 10 April 15, 2008, <http://finance.senate.gov/imo/media/doc/041508dstest.pdf> (assuming such constraints as a probability).

265. ISO 9241-11, *supra* note 12, at § 7.2.2. ISO describes such an approach as follows:

If a work system is judged to be unsatisfactory, systematic analyses of the contribution of different components of the context of use should be conducted. Both direct contributions and the interactions between the components of the context of use should be considered in order to determine the principal causes of the problem. This process may also be used to identify which components are amenable to change, in order to bring about improvements in the overall work system. Diagnostic activity relating to the context of use is often necessary to determine whether problems are due to the product or other components of the work system.

Id.

266. See COMPARATIVE INFORMATION SERIES, *supra* note 26.

267. See *id.*

Still, when usability is the objective, electronic services should not be implemented uncritically; they should be evaluated to determine whether the initiatives really would improve levels of achievement and user comfort – as perceived by users. Some services offered by revenue bodies might be improvements from a compliance cost perspective (or in terms of total resource cost, including the revenue body's cost), but not necessarily from a usability perspective, and might, if made mandatory, actually reduce the level of usability.²⁶⁸

Another option for usability improvements is to assign more tasks to entities with management competencies greater than those of individual taxpayers, such as employers and financial institutions. The example of assigning to charitable associations the duty to submit information about donations exemplifies both utilization of greater management competencies and utilization of information technology.²⁶⁹ Likewise, due to their greater ability to build skills and utilize available technology, professionals can generally master tasks at a higher level of usability than can individual taxpayers. Performance of many tax-related tasks will occur very infrequently for individual taxpayers: annually, or, for some irregularly-recurring tasks, even more rarely.²⁷⁰ Matters, which influence higher or lower levels of usability, such as “time to learn,” “speed of performance,” “rate of errors by users,” and “retention over time”²⁷¹ (retention meaning how well users maintain their knowledge, which is linked to frequency of performance), all lead to an expectation that generally higher levels of usability will be achieved by income-tax professionals than by individual taxpayers.²⁷²

268. See, e.g., *id.* at 185 (referencing electronic filing as benefiting both revenue bodies and taxpayers, and referencing the incidence of mandatory e-filing requirements).

269. See *supra* Part IV.D.1.

270. See, e.g., GRAETZ & SCHENK, *supra* note 1, at 27 (discussing taxable year).

271. SHNEIDERMAN & PLAISANT, *supra* note 12, at 16.

272. Bradford notes that:

[The] three forms of [compliance, transactional and rule] complexity are related, but improving the law with respect to one form may make it worse with respect to the others. Moreover, an arrangement that enhance simplicity of the system for one group of taxpayers may make it more complicated for others. For example, a rule that required employers to prepare the tax return of their employees would simplify the compliance problem of the employees by shifting it to the employers.

BRADFORD, *supra* note 88, at 266-67.

However, as shown by the example of shifting the obligation to submit information regarding charitable contributions, such changes can still be consistently evaluated to

Obtaining usability improvements through accommodations for specific user segments, such as by tailoring contexts of use to better fit their skills and the interplaying environment (for example, by adapting information reporting requirements to the ways some segments of taxpayers can most easily perform these tasks), might also be possible. This approach, however, could imply legal distinctions between segments of taxpayers and consequently entail more differentiated legal authorities for taxation, with the possibility that the tax-law authorities in question, described as products, might become more "complex".²⁷³ But since attributes of legal authorities and assessments of outcome of use are distinct phenomena, as established *supra*,²⁷⁴ changes that give legal authorities more complex attributes might nonetheless be successful ways to improve usability.²⁷⁵

E. *Low or High Usability cf. Complex*

Of course, one could apply the label "complex" generally to outcomes of use that have been assessed through the usability measures and appraised as unacceptable. Likewise, one could define "complexity," when it pertains to taxpayers' encounters, as corresponding to a lower level of usability. Such a definition could embrace the disparity between the user's intended outcome and what actually occurred, the resources spent in relation to the measured effectiveness, and any decrease in the level of users' psychological well-being as a result of the performance. However, applying the "complexity" label in this way does not appear to further elucidate taxpayers' occurrences. Rather, when policymakers' interest lies in what users encounter empirically when performing compliance-related tasks, and the focus of an assessment concerns empirical occurrences beyond mere resource-spending, the terminology of low or high usability level seems to provide the most clarity.²⁷⁶ Furthermore, the usability model's focus on assessing various empirical occurrences highlights the possibility of mitigating undesirable occurrences

determine whether an alteration is an improvement according to a specific objective such as usability or compliance cost. See *supra* Part IV.D.1.

273. See *supra* Part III.B.1.

274. See *supra* Part III.B.

275. See *supra* Part III.B.1.

276. A legal scheme can be pronounced "more usable" in comparison with an alternative scheme. See *supra* IV.D.1. (discussing comparisons between relevant alternatives).

not only through changes in the scope of the tax scheme, but also through changes in the context of use.²⁷⁷

V. ENHANCEMENT OF COMPREHENSION OF WHAT HAPPENS FOR TAXPAYERS

Section A. of this Part argues for the recognition of widespread inadequacy in the ability of current notions of tax complexity to assist in gathering knowledge that is analytically important for possible improvements. It contrasts this with the advantages of applying the usability model. The section then points to the value of more research into assessable aspects of what occurs for users in their performance. Finally, Section B. describes some conceptual confinements of the usability concept, and ends by discussing how computation of tax liability can illustrate several objectives of the field of tax complexity as well as the effect of usability-based analysis.

A. *More Suitable Tools Than Current Notions of Tax Complexity*

The value of a concept for analysis of what occurs for users comes from its capacity to elucidate the *empirical* occurrences comprising taxpayers' outcomes of use – for example, the occurrences that derive from performing activities to obtain a charitable-contribution deduction.²⁷⁸ This is an analysis of processes, and in that respect, “tax complexity” has a number of shortcomings. In general, it is unclear which aspects of outcome of use “tax complexity” embraces, and thus what constitutes “complex” or “not complex”; another problem is that if “tax complexity” is to be comprehended as an estimate, it is unclear how to quantify it.²⁷⁹ For instance, if the empirical occurrences being analyzed pertain to obtaining charitable deductions, which specific occurrences does “tax complexity” address? And if something is “too complex,” then for which specific aspects of taxpayer experience is there an aspiration for improvement? Likewise, if the question is what has caused tax complexity, then what exactly is the thing that has been caused?

277. See for instances, *supra* Part IV.D.1., describing the case with Denmark, which removed the individual taxpayer filing requirement while preserving the right to take a deduction for charitable contributions.

278. See generally *supra* Part I. and III.B. (explaining that empirical occurrences imply welfare propositions).

279. Of course, if “complexity” is merely applied as a synonym for compliance cost, this is not an issue. See *supra* Part II.A.1.

The concepts of compliance cost and psychological cost are clearer, at least conceptually, than ambiguous complexity notions with regard to which aspects each concept embraces and how they enable estimates of what is more or less.²⁸⁰ Although Roberts' definition of tax complexity²⁸¹ does stipulate which aspects it concerns (*i.e.*, achievement and cost), it is vague in the way it conceptualizes when outcome of use warrants the label "complexity."²⁸² Exceeding thresholds of "reasonable" and becoming "excessive" are Roberts' criteria, but by applying these criteria, his complexity terms mix assessment with appraisal of the assessed result.²⁸³

Still, it is instructive to contrast Roberts' definition with usability, because it appears that his real worry can be expressed as low levels of usability in user performance; in this case, tax advisers' performance. Thus, the objectives to which Roberts aspires appear to be high levels of achievement in task performance and acceptable levels of resource-spending, in line with his expectations as a policymaker.²⁸⁴ Evidently, Roberts wants analyses, which can show whether the tax advisers are actually unable to accomplish the intended outcome of a "reasonably certain conclusion," or whether the resources spent are high in relation to the level of achievement, so that in these cases, improvement initiatives can be tailored to mitigate the occurrences that are appraised as not being good enough.²⁸⁵ That kind of analysis requires not only clarity about the occurrences it encompasses – which Roberts' definition has – but also concrete measures, which Roberts' definition lacks but which the usability concept offers in the form of the effectiveness and resource-efficiency measures. By employing such measures, concrete assessed results can be reached, and policymakers can then appraise whether they see results as reflecting a sufficient or insufficient level of usability (once a satisfaction measure has been added).²⁸⁶

280. In practice, the concept of psychological cost has not found any adequate solution for estimation. See *supra* Part II.A.2.

281. See *supra* Part II.B.1.

282. See *supra* Part II.B.1.(applying the indistinct criteria "reasonable" and "excessive" in classifications of tax complexity).

283. See *supra* Part II.B.1.Cooper, cited in note 2 *supra*, is somewhat in line with Roberts, but argues that certain terms occurring in definitions of complexity, such as "reasonable" and "appropriate," show that complexity is what he calls a "relative concept" that can only reflect the outcome of a judgment. See Cooper, *supra* note 2, at 242.

284. See *supra* Part II.B.1.

285. See Roberts *supra* note 2, at 34-36.

286. See *supra* Part IV.A. (describing the ISO 9241-11 definition of usability).

Roberts' definition does not encompass aspects of comfort and discomfort, and thus, in the terminology of usability, concerns only occurrences assessed by the effectiveness and resource-efficiency measures, not users' response to their performance as it is assessed by the satisfaction measure.²⁸⁷ Roberts' definition concerns professional users, and their psychological well-being traditionally has not been a matter of concern.²⁸⁸ Although professionals performing activities for tax-related tasks, either as highly trained tax advisers or as bookkeepers or payroll and clerical staff, would presumably have a professional attitude toward their work, nonetheless it seems contrary to contemporary knowledge about factors influencing productivity to treat their response to the performance as unimportant. It is more realistic to expect that users in the tax context are like users in general,²⁸⁹ and that the psychological well-being of tax professionals in response to their performance of tasks will influence their productivity – and, in turn, the costs of their services. For example, for a user segment of bookkeepers performing activities involved in managing the tasks of businesses' tax accounting, empirical occurrences in the form of stress or anxiety in response to performance is likely to influence productivity. A usability assessment would embrace this aspect of performance through the satisfaction measure, elucidating whether the level of comfort was high or low and thereby providing knowledge important in designing improvement initiatives.²⁹⁰

The issue of taxpayers' unintentional noncompliance is another inquiry which usability could embrace more fully than do references to tax complexity. Unintentional noncompliance is the situation when users' intended outcome is fully lawful compliance, but the actual performance results in an achievement of less than fully lawful compliance.²⁹¹ Thus, unintentional noncompliance directly concerns what occurs with

287. See *supra* Part II.B.1. (discussing Roberts' definition).

288. See, e.g., SANDFORD ET AL., *supra* note 35, at 18 (appearing to make a sharp distinction between "most people," who may experience negative psychological well-being, and, on the other hand, "professionals").

289. See, e.g., JORDAN, *supra* note 201, at 17 (arguing that performing under conditions of low usability causes lower job satisfaction, which has negative effects on productivity); Abran et al., *supra* note 20, at 327 ("For managers, usability is a major decision point in selecting a product, as this decision will have a direct influence on the learnability of the chosen system, and hence on the productivity of those who use it.").

290. See *supra* Part IV.B.3.

291. See generally COMPLIANCE WITH THE TAX RULES BY PRIVATE INDIVIDUALS IN DENMARK, *supra* note 259, at 57.

regard to achievement.²⁹² A usability assessment would elucidate this through the “effectiveness” measure, which in such a situation would determine that the users did not accomplish their intended outcome, causing the performance to display a lower level of usability.²⁹³ Any tax complexity notion would similarly have to include a measure pertaining to users’ empirical level of achievement in order to analyze when unintentional noncompliance is occurring. Whether the legal authorities in question would be described as complex or not – a product description – is irrelevant, given that noncompliance is a process-oriented assessment.²⁹⁴

Inquiries pertaining to what occurs for taxpayers are always about the occurrence of something empirical.²⁹⁵ This also holds for references to “tax complexity” that pertain to what happens for taxpayers.²⁹⁶

292. See *supra* Part IV.B.1. (explaining that the occurrences of achievement concerns what user accomplished through her or his performance).

293. See *supra* Part IV.B.1.

294. See generally *supra* Part III.B. However, if causality were established between a user segment’s less than fully lawful compliance and certain attributes of the legal authorities in question, all within the relevant context of use, this insight might be helpful in the work toward improvement. See *supra* Part III.B.2.

295. See, e.g., *supra* Part III.B, the example. (concerning resource-spending derived from performing activities to obtain a charitable-contribution deduction).

296. A good illustration of how various empirical occurrences are the real phenomena behind the terms “complexity” and “simplicity,” in all their cognates, appears in Heather Field’s article *Binding Choices: Tax Elections & Federal/State Conformity*. (February 14, 2012) (unpublished manuscript) (on file with Houston Business and Tax Law Journal). Referring to explicit tax elections in the federal income tax law, Field discusses the consequences of decisions by states whether or not to bind each taxpayer to the taxpayer’s federal tax choices for state purposes as well. One section of the article bears the heading “Simplicity”:

State conformity to a federal tax election advances simplicity in that a taxpayer only needs to understand one set of rules. . . Consider simplicity of recordkeeping and tax preparation. Even where a state conforms to a federal election, taxpayers’ recordkeeping burdens and tax preparation costs are not simplified if the taxpayer can make a state tax choice that differs from the taxpayer’s federal choice. . . . Similarly, the likelihood of taxpayer mistake is reduced and the likelihood of taxpayer compliance is increased only if taxpayer’s federal tax choice is binding for state tax purpose . . .

Id. at 19.

The underlying empirical occurrences Field refers to in this section are resource-spending and level of achievement. Thus, for Field, “simplified” denotes less resource-spending and a higher level of achievement. Field continues her analysis:

Taxpayers’ ability to take tax issues into account when making business decisions also depends on whether federal elections are binding for state tax purposes. If the taxpayer can make independent choices, then he must analyze which option better reduces his federal income tax and which option better reduces his state income tax. If the taxpayer is required to make the same federal and state choices, then in addition the taxpayer must compare the federal and state tax savings/costs in order to determine which election is tax minimizing, on net. This additional step in the analysis adds complexity,

In the context of analysis, references to tax complexity should not be relegated to the status of a hovering cloud: if anything, they should explicate the concrete phenomena that the term “tax complexity” is meant to include. Amplification both serves the general and commendable aim of clarification and renders comprehensible the idea that the undesirable occurrences might be ameliorated through different mitigation efforts, including changes in context of use, that are tailored to specific encounters appraised as unacceptable.

This Article relies on common designations of aspects of what users encounter, such as time consumption and stress. Research, which goes beyond the conceptual level and aims to include data, would have to expound upon and clarify the occurrences it is to include. Moreover, this Article calls attention to the potential value for further research in exploring whether some aspects of what taxpayers encounter during performance could be characterized as “uniquely complex,” that is, whether some occurrences are distinguished by characteristics such as incidents of choice or contingency; and whether applying to these occurrences a particular definition of complexity adopted from the general complexity literature would make possible a more comprehensive understanding of taxpayers’ experience.²⁹⁷ This is important because such aspects could then be better incorporated into assessments, and possibly mitigated when those assessments reveal levels that, according to policymakers, demonstrate that taxpayers’ performance is not good enough.

Finally, the distinction between descriptions of the attributes of a legal authority (product-oriented) and assessments of occurrences deriving from performance of activities to fulfill tasks induced by the legal authority (process-oriented) highlights the incautiousness of attempting to define tax complexity as embracing both phenomena simultaneously.²⁹⁸ Ascribing “complexity” to a legal authority based on the assessed result of the task performance – i.e., the legal authority is

particular where a specific choice may reduce the taxpayer’s federal tax burden but may increase the taxpayer’s state tax burden (or vice versa).

Id. at 20.

Here, “complexity” seems primarily to comprise resource-spending, but also includes occurrences in which a taxpayer experiences incidents resulting from choice and contingency. See *supra* Part III.C.4. (discussing choice and contingency as conceptions of complexity).

297. See Figure 4, *supra* Part III.D.4., and accompanying note 171 (referring to hypercomplexity as a definition of complexity, which has not been employed to describe taxpayers’ experience).

298. See *supra* Part III.B.

complex if the performance is appraised as unsatisfactory, and simple if the performance is satisfactory – would overlook the crucial point that what happens for taxpayers (their outcome of use) is not intrinsic to the legal authority.²⁹⁹ Were a definition to attempt to embrace both phenomena, the designation of a specific legal authority as either complex or simple would not be fixed but, rather, would fluctuate depending on the context of use in the particular assessment.

B. *Confinements by Objectives: Computation of Tax Liability as Illustration for Summing Up*

Each of the concepts discussed has natural limitations, based upon its core objectives: for example, compliance-cost studies pertain to elucidation of economic cost, while the usability concept pertains to “quality in use.” If policymakers’ interest concerns different objectives than these, neither the compliance-cost concept nor the usability concept would be adequate models for analyses.

“Political legitimacy,” often discussed in the field of tax complexity and simplification,³⁰⁰ is an objective quite distinct from lower compliance cost or higher usability.³⁰¹ The core of this discussion is the assertion that “tax complexity” undermines political legitimacy because it is contrary to good governance.³⁰² To what degree the literature concerning political legitimacy and tax complexity has convincingly defined what is meant by political legitimacy – and especially which indicators of citizens’ behavior are adequate to assess the strength of political legitimacy, and how that behavior is influenced by the real phenomena behind the “tax complexity” designation³⁰³ – is a discussion beyond the aims of this Article. Whether political

299. See *supra* Part III.B.

300. See, e.g., Cooper, *supra* note 2, at 244 (discussing, with reservations, the scholarly merits of the branch of tax literature concerning political legitimacy when it is amalgamated with “plain English” claims).

301. See *supra* Part II.A. (explaining that the compliance cost concept assesses resource-spending understood as economic cost); see also *supra* Part IV. (explaining that usability assesses “quality in use”).

302. See, e.g., Roberts, *supra* note 2, at 32 (“It surely should not be necessary to prove that a democratic system of government requires that its citizens know the law they are enjoined to obey.”).

303. To the extent that the literature would claim causality between taxpayers’ occurrences deriving from their performance of tax-compliance tasks and the relevant indicators of political legitimacy, elucidation of whether the causality would relate to taxpayers’ factual encounters or their perceptions of what occurred would be important knowledge. See *supra* Part IV.C. (describing assessment of what happened factually as opposed to what happened accordingly to users’ perception).

legitimacy is actually strengthened by generally high levels of usability in managing tasks induced by a tax system, including instances in which high usability is attained primarily through changes in contexts of use, is also a discussion beyond the scope of this Article. As noted, the usability concept itself is obviously not a way to assess levels of political legitimacy.³⁰⁴

An essay by Lawrence Zelenak, *Complex Tax Legislation in the TurboTax Era*, is a recent example of work in the field of tax complexity with political legitimacy as the stated objective of its discussions and recommendations.³⁰⁵ The essay's content regarding complex tax legislation and the prevalent use of tax software is also an excellent illustration of the usability framework and the usability concept's applicability; for that reason, reference to that essay provides a good way to sum up the central part of this Article.

Zelenak argues that at the time when all tax returns were prepared by hand, Congress was greatly constrained in its ability to impose what he calls "computationally complex provisions" on large numbers of taxpayers.³⁰⁶ But since computer software today has revolutionized the preparation of federal income tax returns, this is not any longer the case.³⁰⁷ The constraint came from the fact that taxpayers would have objected vociferously to complex provisions, due to the computational burden they would have had to bear or, alternatively, the price demanded by paid preparers to do the work.³⁰⁸ Today this burden has been alleviated by the widespread adoption of tax software.³⁰⁹ No longer constrained, Zelenak argues, "Congress has responded by imposing unprecedented computational complexity on large numbers of taxpayers – primarily through . . . the alternative minimum tax and the proliferation of phase outs of credits, deductions, and exclusions."³¹⁰ As taxpayers and their tax preparers have no problems doing the computations in practice due to the software, Zelenak opines that this state of affairs would not be problematic were it not that the inner workings of the computationally complex provisions are beyond taxpayers' comprehension and that "[t]his undermines both the political legitimacy of the tax

304. See *supra* Part IV. (explaining that usability assesses "quality in use").

305. Zelenak, *supra* note 26, at 91.

306. *Id.* at 92.

307. *Id.*

308. *Id.*

309. *Id.*

310. *Id.* at 91, 98-99.

systems and the ability of taxpayers to engage in informed tax planning.”³¹¹

As noted, the content of Zelenak’s essay is a good illustration of the applicability of the usability framework to create clarity in the field of tax complexity. In the terminology of this Article, Zelenak is describing how the legal authorities governing the computation of U.S. taxpayers’ tax liability have become more complex, for example as they relate to the proliferation of phase-outs.³¹² This is a product-oriented description that focuses on characterizing the attributes of legal authorities by applying a general understanding of complexity as rising in parallel with an increasing number of elements, distinctions, and possible relations among the attributes of the described objects.³¹³ Reasoning along these lines, Zelenak reaches the conclusion that the legal provisions involving computations display an increasing level of complexity, which he calls “computational complexity.”³¹⁴ However, Zelenak also makes it clear that, to use the terminology of this Article, the taxpayer’s outcome of use has not deteriorated, and thus that a process-oriented assessment, for example as compliance cost or as usability, would show that the development of increased complexity in the legal authorities has not caused problems in terms of such objectives.³¹⁵ The observation that outcome of use has not deteriorated in spite of increased complexity in the product, i.e., the legal authorities in question, illustrates the fact that outcome of use is not intrinsic to the product, but is influenced by the context of use, which in this case has changed simultaneously as tax software has replaced pencil and paper.³¹⁶ Hence, Zelenak’s essay³¹⁷ reinforces the point that taxpayers, who are subject to more complex legal authorities, might nonetheless have an outcome of use reflecting the same or even an improved level.³¹⁸

If complex provisions governing the computation of tax liability had been imposed before the availability of tax software, or even mass-market financial calculators, the burden taxpayers

311. *Id.* at 91.

312. *See id.* at 98-99.

313. *See supra* Part III.B.1.

314. *See Zelenak, supra* note 26, at 98.

315. *See id.* at 92.

316. *See supra* Part III.B.

317. Zelenak establishes that although the legal authorities have risen in complexity, there, nonetheless, has been no increase in the taxpayers’ burden, and that this phenomenon is due to the shift in the way taxpayers deal with compliance tasks, i.e., by using tax-preparation software. Zelenak, *supra* note 26, at 91-92.

318. *See supra* Part III.B.1.

would have borne as a result would have manifested itself in the form of occurrences deriving from computational activities, such as extreme time consumption, lack of achievement due to errors or giving up, and great frustration as a response to performance.³¹⁹ These are the empirical occurrences, which a usability evaluation can assess through its three measures; and in such a scenario, a usability assessment would have shown a low level of usability.³²⁰ Zelenak's other objective, the ability of taxpayers to engage in informed tax planning,³²¹ actually relates to taxpayers' performance of an intended task and can also be analyzed from a usability point of view. According to Zelenak, the current U.S. provisions impede taxpayers' ability to engage in well-informed basic tax planning, or to respond appropriately to the many incentives Congress has embedded in the tax law.³²² Whether this is the case for taxpayers whose intended outcome is to successfully utilize such options, and who therefore undertake those compliance tasks, can be analyzed by employing the three usability measures. The measures can assess the occurrences brought about by carrying out the compliance-related activities: the effectiveness measure, assessing the degree to which a user segment achieved the attempted tax planning; the resource-efficiency measure, assessing the resources spent in relation to the achieved goals; and the satisfaction measure, assessing users' responses to performing the activities, in the form of stress and the like.³²³ Zelenak's proposition, which is that taxpayers are currently unable to engage effectively in informed tax planning,³²⁴ could be documented by these measures' demonstration of a low level of usability.

As with other attempts to improve usability, amelioration may be possible not only through changes in the tax scheme, which Zelenak advocates, but also through changes in the context of use.³²⁵ For instance, new tax software could be developed which would enable a taxpayer to succeed in well-informed basic tax planning with less time consumption, such as by determining her or his marginal tax rate and the after-tax cost of charitable

319. See Zelenak, *supra* note 26, at 92 (concerning computational burden).

320. See *supra* Part IV.B. (discussing the three usability measures and their application to assessments of time consumption, level of achievement, and frustration).

321. See Zelenak, *supra* note 26, at 103.

322. *Id.*

323. See *supra* Part IV.B.

324. See Zelenak, *supra* note 26, at 103.

325. See *supra* Part IV.D.2.

giving at the appropriate time so that the taxpayer can decide whether to contribute.

What Zelenak calls the TurboTax Era illuminates one more point relevant to this Article. Taxpayers seek to be able to handle their tax affairs with the level of usability that the newly available software systems provide.³²⁶ In other words, taxpayers want usability,³²⁷ and actual practice has for a long time had usability as its aim.³²⁸

VI. CONCLUSION

This Article has suggested a valuable new concept, "usability," for assessment and analysis of the experiences of taxpayers and others who do work on their behalf in performing tax-compliance tasks. In addition, the Article has examined and critiqued existing tax complexity notions and the concepts of compliance cost and psychological cost. Taken together, these aspects of the Article provide an in-depth understanding of available ways to assess what happens for taxpayers as a consequence of performing tax compliance activities.

The usability model demonstrates that assessment of taxpayers' experiences in performing compliance tasks requires drawing up principles from the study of human-product interactions. Consequently, assessments must take into account how different taxpayers actually carry out their compliance tasks – such as, most prominently, how the pervasive use of tax-preparation software has noticeably eased the performance of many compliance-related tasks. By contrast, a mere description of attributes of the tax rules themselves, such as the number of articles and their substantive interconnections – isolated from how taxpayers actually carry out their tasks – is a particular kind of study, and does not directly concern occurrences that taxpayers might encounter. For instance, taxpayers who are

326. See generally Bankman, *supra* note 52, at 1432 ("More than 98 percent of those who used the ReadyReturn said they would use it again.").

327. Schenk has succinctly made a comparable point: "The man on the street is concerned with the complexity that manifests itself each year on April 15, when he tries to do his tax return . . . simplification means those steps that would ease this annual burden." Deborah Schenk, *A Simplification for the Average Taxpayer*, in CHARLES H. GUSTAFSON, ED., *FEDERAL INCOME TAX SIMPLIFICATION* 115, 115 (1979).

328. Concerning a concept's possible development periods for reaching maturity, see Pope's description, *Tax Simplification*, *supra* note 34, explaining how compliance-cost studies reached the prominence they have today only gradually. Acceptance of usability would not have to go through the same gradual development, but could skip some periods because the existing usability concept, developed outside the legal context, is already available to be built upon, and because tax practice has already implicitly or explicitly recognized the usability objectives.

subject to extensive rules for computation of tax liability might nonetheless encounter acceptable levels of resource-spending and frustration. The current literature's frequent references to "tax complexity," without thorough explication of the phenomena the term is meant to embrace, obscures that fact. This Article recommends that if the term "tax complexity" is intended to refer to taxpayers' experiences, then the discussion should always clarify what the specific concerns are.

What follows is a note of optimism in the field of tax complexity and simplification. The usability framework illustrates that increasing tax complexity, understood as the confluence of tax statutes and other legal authorities, should not automatically be cause for despair. Numerous initiatives related to taxpayers' actual performance of tasks, such as increased application of information technology, can improve the resulting levels of resource-spending, achievement, and frustration despite increases in the complexity of the tax rules. Moreover, the usability concept's more comprehensive assessment of taxpayers' empirical occurrences makes a significant contribution to the field. The usability concept accentuates the critical importance of whether taxpayers in fact succeed in accomplishing their compliance tasks; it proposes a way to actually measure psychological well-being; and it emphasizes that inquiries into what happens for taxpayers can examine both how they performed as an objective matter, and how they themselves perceived their compliance efforts. All of these considerations have important welfare implications.

Analyses based on the usability model can enhance policymakers' ability to tailor proposed changes more precisely toward those taxpayer occurrences in need of improvement. And the guiding principle to aim for is simply that users in the income-tax context desire smoothness in their performance, just as they do in the other activities in which they engage.