

A genealogy of the scalable subject: Measuring health in the Cornell Study of Occupational Retirement (1950–60)

History of the Human Sciences
2023, Vol. 36(2) 128–153
© The Author(s) 2022



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/09526951221113438
journals.sagepub.com/home/hhs



Tiago Moreira  ID
Durham University, UK

Abstract

Increased use of scales in data-driven consumer digital platforms and the management of organisations has led to greater interest in understanding social and psychological measurement expertise and techniques as historically constituted ‘technologies of power’ in the making of what Stark has labelled the ‘scalable subject’. Taking a genealogical approach, and drawing on published and archival data, this article focuses on self-rated health, a scale widely used in population censuses, national health surveys, patient-reported outcome measurement tools, and a variety of digital apps. The article suggests that the first methodological articulation of self-rated health by the investigators of the Cornell Study of Occupational Retirement (1951–58) provides a window into the key epistemic, institutional, and cultural uncertainties about psychological and social measurement, processes of adjustment to ‘old age’, and the capacity of individuals to value their own health. I propose that these uncertainties have become incorporated into extant and operational measurements of health.

Keywords

genealogy, Guttman, politics of method, psychometrics, self-rated health

Corresponding author:

Tiago Moreira, Durham University, Department of Sociology, 32 Old Elvet, Durham, DH1 3HN, UK.
Email: tiago.moreira@durham.ac.uk

Introduction

Contemporary widespread use of scales in data-driven consumer digital platforms and the management of organisations has stimulated increased interest in measurement and quantification in the humanities and social sciences. Historical research has detailed how, from their development in the late 19th century by Galton to their establishment as routine forms of assessment in institutions such as schools or workplaces in the interwar years (e.g. Carson, 2007), psychometrics became key to the management of organisations in modern societies. By enabling the mapping of individual attributes in mathematical, structured measurement, psychometrics facilitated the generation of inscriptions – distribution curves, and so on – that can be easily manipulated, and its results can be readily compared, correlated, subsumed, integrated, and understood in graphical form, such as position on a normal curve. Indeed, the power of representation and technique was so important in the consolidation of psychometrics that Danziger claimed that by the turn of the 1940s ‘investigators, in their almost exclusive reliance on tests, seemed to have substituted technology for science’ (Danziger, 1990: 165). In this process, the embedding of applied psychology techniques within modern, bureaucratic institutions in the interwar period, particularly in the US, aimed to facilitate and guide the *adjustment* of individuals to new social and technological environments (e.g. Schilling and Casper, 2015).

Understanding the formation and shifts that led to the establishment of such relationships has also stimulated historians’ interest in the role of data collection and processing technologies such as questionnaires (e.g. Igo, 2007; Young, 2017). A key insight from such research is that surveys and social and psychological measurement are performative, and not merely descriptive of the phenomena they study. In line with this, Desrosières’ work calls our empirical attention to the establishment of ‘conventions of equivalence’ underpinned by a series of ‘comparisons, negotiations, compromises, translations, inscriptions, codings, of codified and replicable procedures and calculations’ (Desrosières, 2008: 11; my translation) that bring collective entities such as ‘attitudes’ into existence (see also Desrosières, 1990). Thus, Desrosières proposes that it is in becoming transferrable, combinable, and routinised that such forms of measurement become performative, being able to effectively reconfigure practices, norms, and institutional arrangements. Understanding the transformative capacity of psychometric procedures relies on carefully and empirically tracing how fuzzy entities such as ‘experience’ or ‘attitude’ can become tractable, measurable, and actionable within specific contexts and settings. My approach to this in this article is particularly sensitive to how controversy and uncertainty is generative of specific mensuration practices and associated institutional arrangements: how analysing *the politics of method* enables us to understand society-in-the-making, providing insight into ‘what allows the social to be established’ (Latour, 2005: 25).

This article focuses on health measurement. This is significantly motivated by the increased role played by digital monitoring and tracking technologies in making health visible to mundane activities in contemporary societies (e.g. Lupton, 2018), and by how these are institutionally linked to wider processes of ‘datafication’ within national health systems (Hoeyer, Bauer, and Pickersgill, 2019). Of the wide variety of health measurement tools deployed in these processes, self-rated health (SRH) stands out for

its significance, durability, and bearing. Versions of the SRH scale have been used since at least the 1950s in settings such as the British General Household Survey, the US National Health Interview Survey, and health surveys in most OECD countries (Bowling, 2005). It has been included in the UK census questionnaire since 2001 (Office for National Statistics, 2013), and is a regular item in patient-reported outcome measurement (PROM) tools and health apps. SRH is normally defined as an individual's subjective assessment/perception of their own general health status, its items consisting of respondents' own classification of their health as 'excellent, good, fair, poor or very poor'. Such self-classification is statistically predictive of morbidity and mortality (Falconer and Quesnel-Vallée, 2017; Idler and Benyamini, 1997), making the cognitive and affective processes that underpin responses an important topic of research (e.g. Jylhä, 2009).

This article traces the development of the measurement of SRH as a 'history of the present' of health, and details the epistemic, technological, and institutional 'heterogeneity of what was imagined consistent with itself' (Foucault, 1991: 82). The article can thus be said to be a possible genealogy of what Stark (2018) has labelled the 'scalable subject', in that it outlines the myriad, partial interferences that brought about the contemporary recombination of *computational* and *psychological* sciences encapsulated in the emergence of contemporary digitally mediated subjectivity.

The received history of SRH, mainly manifest in methodological writings, portrays it as a linear stabilisation process culminating in 1992 in the Medical Outcomes Study's 36-Item Short-Form Health Survey (SF-36) questionnaire (Ware and Sherbourne, 1992), and presents psychometrics as a key technique in its development and validation (Bowling, 2017; McDowell, 2006). Another version of its history links its delineation to debates within social gerontology in the 1950s and 1960s (Tissue, 1972), and particularly to Maddox's (1962, 1987) rethinking of the health items included in the 'Cavan Inventory' of attitudes and activities generated as part of the Personal Adjustment in Old Age study (PAOA, 1943–9; Burgess *et al.*, 1949). Although more embedded in the uncertainties experienced by researchers in attempting to establish non-clinical measures of health, this account is still reliant on the methodological power of the psychometric scaling procedures proposed by Thurstone (1927) – as deployed by Cavan in the PAOA study – for the generation, selection, and weighting of items in a questionnaire to produce a reliable measure of health.

As recognised by Tissue (1972; see also Ware, Davis-Avery, and Donald, 1978), the first methodological articulation of SRH was formulated by the investigators of the Cornell Study of Occupational Retirement (CSOR, 1951–58) in a paper on the 'validity of health questionnaires' (Suchman, Phillips, and Streib, 1958). The main reason for focusing on the activities of the CSOR is not, however, related to historical precedence, but instead linked to how the CSOR provides a window into the epistemic, methodological, institutional, and cultural dynamics through which SRH was developed. Not usually included in the list of longitudinal studies of ageing instigated in the US in the 1950s (Achenbaum, 1995; Moreira, 2017), the CSOR is unique in its use of the techniques developed by Louis Guttman (1944) for scaling attributes, and of a Columbia Sociology-style methodological approach to the issues of retirement, social adjustment, and health (Pienta and Lyle, 2018).

Drawing on archival and documentary data, this article argues that the articulation of SRH was encased in three layers of uncertainty, two of them formalised in the shape of controversies.¹ The first section focuses on the challenge set by Guttman scaling to established, psychometric ‘Thurstonian’ methods within psychology and sociology. It also documents the reasons why Guttman scaling became embedded in the activities of the Research Branch of the Morale Services Division of the Army Service Forces during and immediately after World War II, coming to shape the armed forces’ techniques of behaviour management and adjustment of personnel to stress, fear, and similar phenomena. The article then analyses the scientific dispute concerning the measurement of ‘adjustment’ to retirement, an issue widely and publicly viewed as a challenge to the American economy and its modern ‘pragmatic culture’ from the mid 1930s onwards (Achenbaum, 1995; Frank, 1939). This section emphasises the *politics of method* separating the Thurstonian scale of adjustment to retirement proposed by Ernest Burgess and his Chicago collaborators in the PAOA study, on the one hand, and the Guttmanian, configurational approach favoured by the CSOR team, on the other. The final section examines how this politics of method shaped social and psychological researchers’ ability to build non-clinical health measurement techniques. It details the trials and tribulations of cognitively and politically framing and obtaining health as an epistemic object within the social sciences. In the conclusion, I will suggest that the uncertainty thus embedded in the measurement of health has become inherent in the use of SRH in the present day.

The scaling controversy

In the history of social and psychological measurement, the transposition of practices and creation of instruments to be deployed beyond the walls of the laboratory is of crucial importance (Danziger, 1997). This requires a variety of ‘negotiations, compromises, translations, inscriptions, codings’ that modify not only the empirical referent but also the means through which such objects are captured, measured, compared, and so on. Thus, it is possible to identify a gradual process in which the questionnaire as a technique is progressively applied to bureaucratic institutions such as schools, firms, and the army. In developing methodologies to quantify what until then could be described only as fleeting, thin, and fragmented objects, Thurstone’s work on what he labelled ‘subjective measurement’ in the late 1920s is recognised as a fundamental milestone in the history of psychological research.

Supported by the Laura Spelman Rockefeller Memorial, Thurstone’s work at the University of Chicago’s Social Science Research Committee was concerned specifically with developing methods for ‘community based research’. As Bulmer (1980) documents, the committee’s work was guided by the strong position of the Chicago Sociology Department, and in particular Ernest Burgess’ vision that ‘the survey provides a unique opportunity both for investigation and for social construction [where] the analysis of *mental attitudes* [is used] for the study of the control of forces in securing improvement’ (Burgess, 1916: 499; emphasis added).

It is thus no surprise that Thurstone’s key publication on the measurement of ‘mental attitudes’ was first published in the department’s institutional journal, the *American*

Journal of Sociology (Thurstone, 1928; see also Abbot, 1999). This paper was crucial because, in it, Thurstone proposed that the ability to measure attitudes was predicated on their being – ontologically – similar to other complex qualities gauged by single indexes.

Departing from the premise that ‘the very idea of measurement implies a linear continuum of some sort such as length, price, volume, weight, age’, Thurstone (1928: 534) argued that for attitudes to be measured, they had to be conceived as a range of opinions on specific matters such as those that arose around ‘disputed social issues’, from which a ‘scale of evenly graduated opinions’ (ibid.: 554) could be constructed. To do this, Thurstone drew on previous theoretical work where he had proposed that by means of a series of responses to statements it should be possible to build a range of opinions and to locate, with some certainty, an individual’s position within that range (ibid.: 548). This elegant theory was dependent, as Thurstone was aware, on psychological scales being normally distributed within what ‘we shall call the psychological continuum’ (Thurstone, 1927: 273). As a result, Thurstone admitted, ‘the psychological scale [could only be thought of as] at best an artificial construct’ (ibid.: 275).

This theoretical assumption of there being ‘an infinite number of attitudes that might be represented along the attitude scale’ (Thurstone, 1928: 537), which could be gauged through questions as stimuli, was also important because it enabled Thurstone to propose that a ‘more arbitrary’ unit could be used by dividing the scale into 10 equal measures (ibid.: 553). Presented as a compromise between the cumbersome deployment of the law of comparative judgement outside the laboratory and the need to build instruments to measure attitudes on issues of contemporaneous concern (divorce, ‘the Negro’), what came to be known as the Thurstonian method of scale construction was a series of procedures for generating, selecting, and weighting statements ‘on the issue in question’, then asking a sample of individuals to agree/disagree with the resulting 20 statements, the final ‘score for each person [being] the average scale value of all the statements that he [*sic*] has indorsed [*sic*’ (ibid.: 553).

Thurstone’s influence in psychological research, education, and social measurement is well documented and not within the concerns of this article. However, it is important to note that perhaps Thurstone’s most significant route of impact in the social sciences was through the sway held by Samuel Stouffer, starting with his application of Thurstone’s method of ‘equal appearing intervals’ (see above) to the controversial issue of prohibition in the late 1920s and 1930s (Ryan, 2010: 100–6). A zealous advocate of the use of survey methods to study opinions and attitudes in American society, Stouffer played, through his work on various committees and as a leading researcher of social statistics in the Chicago Department of Sociology, a key role in establishing methodological standards in research seeking to guide social policy during the Great Depression. By the latter half of the 1930s, Stouffer had become a major figure in American survey research, arguing that ‘behind any successful study ... stands the mathematical statistician’ (Stouffer, 1941: 58). It may have been this stance, as well as his prominent position at the boundary between academy and policy, that led a graduate student in sociology at the University of Minnesota, Louis Guttman, to seek Stouffer’s help as a mentor in social statistics.

As a fellow of the Social Science Research Council, Guttman was enticed by Stouffer’s pragmatic but robust approach to questionnaire design and application.

He aimed to apply this approach to the sociological problem of measuring social status, on which he was working with his advisor Stuart Chapin (Arbel, 2016; Guttman, 1942). For this, he focused on the statistical techniques Thurstone (1935) had been developing for determining underlying constructs within a set of observed variables. In this investigation, he was struck by how the cumbersome process of selecting and weighting items in the construction of a scale – as proposed by Thurstone (above) – could still result in the ‘arbitrary’ inclusion/exclusion of items in a social status construct (Guttman, 1942: 368).

With Stouffer’s support, Guttman developed an alternative technique whereby qualitative data could be recorded in a manner amenable to treatment by matrix algebra. In the statistical appendix to the Social Science Research Council monograph on the *Prediction of Personal Adjustment*, Guttman proposed that a method could be devised whereby ‘the behaviour of an individual is not considered to be a single value but a distribution of the values of the acts he [*sic*] performs’. He suggested that by retaining the ‘configuration’ of their acts in a ‘class of behaviour’, it would be possible to test ‘the utility of the set of items to predict both the individual’s behaviour when performing other acts and that other populations in the same domain of practice (Guttman, 1941: 323–4). While acknowledging Thurstone’s ‘trail-blazing work’, Guttman (*ibid.*: 345) was effectively proposing a method of scale construction that did away with the requirement to postulate a ‘psychological continuum’ where the opinion of an individual would obtain a single value. The method asked whether there was a *pattern* to a configuration of acts and, ultimately, aimed to provide a way of determining the underlying unity – and ‘reality’ – of the concept being measured.

Guttman was reluctant to emphasise this crucial ontological and methodological difference between his method and Thurstone’s. Describing his approach to Stouffer in 1942, Guttman argued that ‘it would once and for all do away with weighting problems’ and ‘form a rapid, efficient, theoretically sound, and quite easily understandable method of scale construction’ (Guttman to Stouffer, in Ryan, 2010: 183). It solved practical problems in scale construction. Instead of requiring what he referred to as ‘proceduralist’ techniques of item selection and weighting (Guttman, 1944: 141), scaling could be turned into an empirical inquiry. The weights attributed to items resulted from their position in the configuration, and not from arbitrary decisions by ‘judges’, their inclusion resulting from how well they fitted in the configuration. Scale construction was, therefore, an analytical procedure: ‘Scaling analysis is a formal analysis, and hence applies to any universe of qualitative data of any science, obtained by any manner of observation’ (*ibid.*: 142).

It was this methodological versatility, as well as its sound statistical underpinning, that motivated Stouffer to invite Guttman to work for the Research Branch of the Morale Services Division of the Army Service Forces, where the pair, along with Edward Suchman, Paul Lazarsfeld, John Clausen, and others, collected and analysed the data for what came to be known as *The American Soldier*. This setting would provide the ideal context to strengthen the robustness of Guttman’s technique – scalogram analysis – through what Lutz (1997) labels the ‘epistemology of the bunker’ (see also Converse, 1987). An example of the power of his technique was Guttman’s work in developing a predictive scale of fear in battle. Drawing on survey data from 1944, Guttman and colleagues were able to demonstrate that there was a scale pattern to

physiological manifestations of fear, enabling the prediction of individuals' responses within and outwith the scale, and facilitating selection and personnel decisions (Stouffer *et al.*, 1949: 201, Table 3).

The capacity of scale analysis as a technique to establish what Desrosières would label a 'convention of equivalence' was epitomised by the coefficient of reproducibility, a metric of the approximation of scales to 'perfect' rank order, and assisted by a mechanical device designed by Guttman with Suchman's assistance, the scalogram board (see Figures 1–2). On this wooden board, respondents' answers were logged through metal shots in holes. By shifting the slats, the board could be physically manipulated to reveal a scale pattern if one existed, making it visually evident 'at a glance' what items should and should not be included in the scale. The scalogram board was an effective 'immutable mobile' (Latour, 1990), supporting a means of producing and applying results 'which require[s] no knowledge of statistics' (Guttman, 1944: 139).

In predicting adjustment to combat, the scalogram technique was key in assisting the collection and computation of the data collected from armed forces personnel during the World War II conflict, making the preliminary choice, validation, and testing of questions, items, and inventories included in questionnaires of less importance than the analysis of their psychological or sociological 'meaning'. It supported a pragmatic approach to data – what in contemporary terms might be called a data-centric approach – to support the 'engineering mission' of the Research Branch (Stouffer, 1950: v).

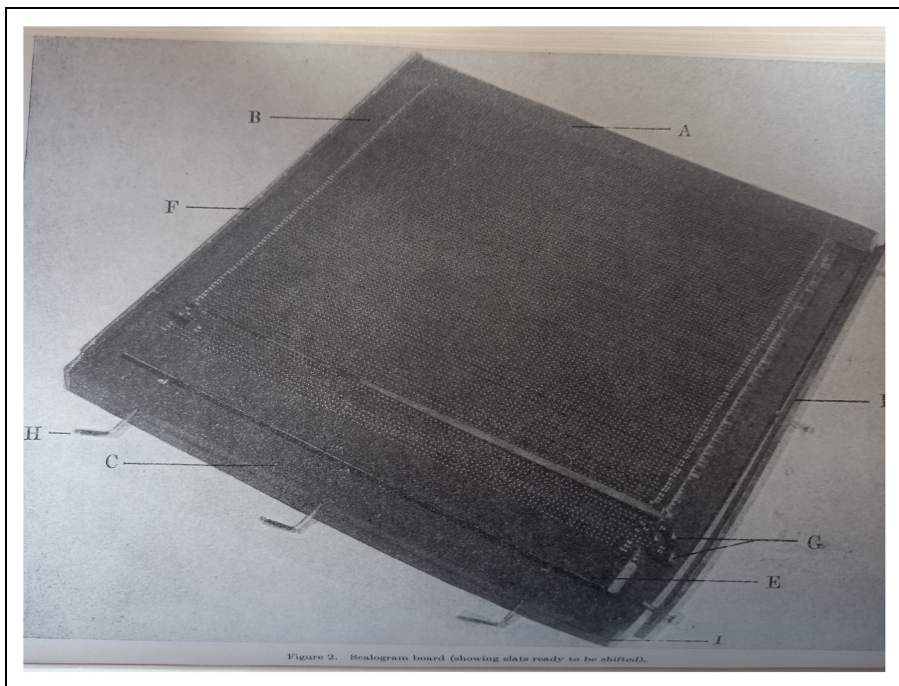


Figure 1. Scalogram board (Stouffer *et al.*, 1950: 92).

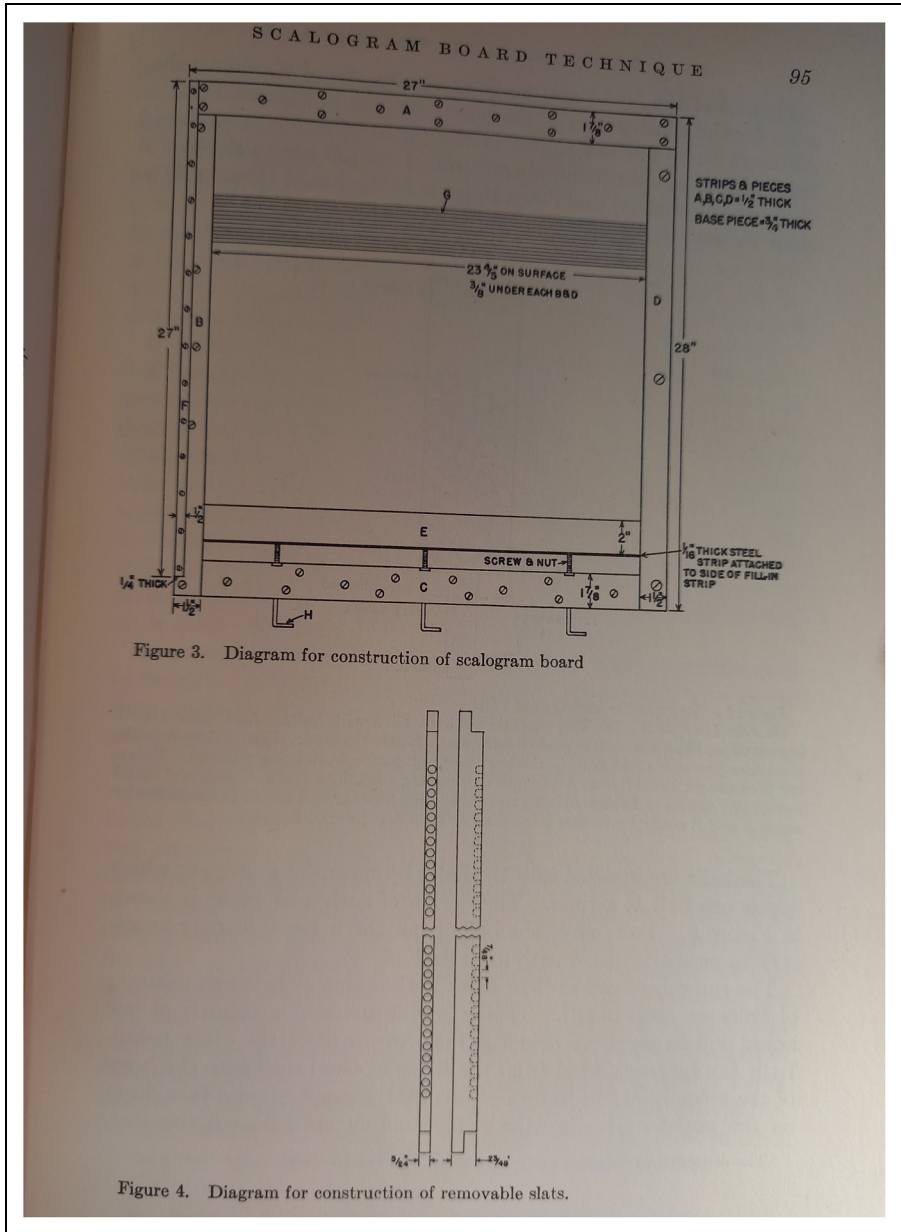


Figure 3. Diagram for construction of scalogram board

Figure 4. Diagram for construction of removable slats.

Figure 2. Diagrams for scalogram board (Stouffer et al., 1950: 95).

This data processing work buttressed the methodological confidence around Guttman's technique, and shortly after the end of the war, a variety of publications outlined the approach and its ontological underpinning (e.g. Guttman, 1947). Thus, recognising the possibility of the debate hinging on 'metaphysical faith in a particular model', Stouffer (1950: vii) was clear in pointing out that the Guttman technique was 'controversial' because it 'dispenses with the concept of [an] underlying continuum to which the response to a particular item is to be relatable' (ibid.). Instead of constructing scales to fit the requirements of statistical theory, hinging on a theorisation of the structure of political disagreement (see above), researchers could use scale analysis to discover different varieties of scales, from linear rank order to U-shaped curves (Guttman and Suchman, 1947) to quasi-scales. In this respect, Guttman's technique could be proposed as one that promised to be a universal data analysis tool underpinned by a different ontology of measurement, which could be used 'as a test of the meaning of items in an effort to eliminate items which do not belong to the scalable universe' (Suchman, 1950: 90; emphasis added).

Again, it is beyond the scope of this article to trace the impact of Guttman scaling in the social sciences. It is important to note, however, that by the mid 1950s and 1960s it had become a staple, basic technique of scaling, particularly within US sociology (e.g. Riley, 1954; Suchman and Francis, 1954). Writing in the mid 1970s in a sourcebook of scaling techniques for social scientists, Maranell claimed that Guttman's technique had 'served to define what is meant by scaling for many people, because it is the scaling method most typically presented and described in introductory methods books' (Maranell, 1974: 129). In parallel, a series of critiques of the method had also emerged, particularly concerned with the possibility that scales might be chance findings, there being 'no definite [statistical] proof that all the items in a given [Guttman] scale are measures of the same dimension' (Schooler, 1968: 296).

The Cornell Retirement Study

It is generally agreed that the publication of Cowdry's *Problems of Ageing* in 1939 marks the consolidation of gerontology as a field of research (Achenbaum, 1995; Katz, 1996; Park, 2008). Drawing together a variety of hitherto loosely connected strands of research on 'old people', the 1937 Woods Hole Conference that gave origin to the book had been instigated by Cowdry's and the Macy Foundation's concern with the effects of the Social Security Act of 1935 in the context of demographic ageing and the Great Depression. In its efforts to define the challenges of ageing for American society, the foundation's own view on the role of social research, outlined by Lawrence Frank, its director, was pivotal. Frank, a sociologist by training, was concerned with how social research could provide the basis for institutional and normative support – what he labelled 'designs for living' – for individuals' adjustment and 'guiding conceptions of life' amid rapid social change (Moreira, 2017: 22–5).

In outlining how misalignment between biological, psychological, and social processes was at the basis of the 'problems of aging', Frank (1946) was actively and explicitly aligning the new field of gerontology with an emerging research agenda: the question of social adjustment. A few years before, in 1941, the Committee on Social

Adjustment of the Social Science Research Council, led by Burgess, had selected adjustment to old age as a field that required active attention. Taking adjustment in 'its common sense meaning [of comprising] all efforts of human beings to find more satisfactory ways of getting along with one another', Pollack, in his Social Science Research Council report a few years later, suggested that research should focus 'on the types of adjustive behaviour which may lead to the solution' of the problem of old age (Pollack, 1948: 38).

Furthering this agenda was Burgess' own study with Havighurst and Cavan on 'personal adjustment in old age' (the PAOA study). Suggesting that changes in attitudes were 'especially important for personal adjustment', especially in a 'rapidly changing society', the study aimed to 'determine the conditions under which changes of attitudes can be brought about' after retirement (Burgess *et al.*, 1949: 14). The study brought together the psychometric expertise of Havighurst, Cavan's technical mastery of both case and statistical methods, and Burgess' own interest – articulated while he was serving as president of the Committee on Social Adjustment – in developing a 'scale for measuring successful adjustment, an essential [tool] for determining how personality and social background are related to adjustment in old age' (Burgess, in Young, 1941: 884). In the field, however, Havighurst's and Cavan's contributions were key to developing such a scale.

Havighurst, who had taken over the directorship of the child and adolescent development programme at Rockefeller after Frank, was especially interested in the development of measures of individual development and personality that were independent of cultural and social assumptions and expectations. This made the domain of 'old age' of particular significance because of the 'methodological problems relating to the technics for studying individuals of widely varying ages' (Havighurst, Kuhlén, and McGuire, 1947: 344). Cavan, the go-to but often unacknowledged researcher in Burgess' many projects since the turn of the 1930s (see Burgess, 1934), was not only very experienced in developing scales within questionnaires but was also an outstandingly methodical data collector and analyst. Her own substantive academic concern with suicide as a crisis of 'adjustment ... in the reciprocal relationship of subjective interest and external world [where the individual becomes] personally disorganised' (Cavan, 1928: 147) was particularly relevant. In developing Burgess' proposed 'scale of successful adjustment', Cavan and Havighurst combined their expertise to identify the factors that drove changes in old age.

In so doing, they explicitly deployed Thurstonian procedures. First, Cavan and colleagues compiled a 'list of attitudinal statements ... obtained from book and articles ... and a number of personal interviews' (Burgess *et al.*, 1949: 112). Then, 'eight judges were asked to give a numerical rank to the statement[s] in each category', ranks that were then analysed and reduced, and a new list was subjected to rating by another set of 21 'mature judges', as well as a group of 27 graduate students in social statistics, and further reduced on the basis of overlap and/or possible misunderstanding. 'Weights were then assigned to [statements] by retaining rank order' (*ibid.*: 113), and the score obtained by calculating the sum of the scores in the 10 categories of statements, such that the higher the score, 'the more adequate was the individual's adjustment'. This was subjected to internal consistency analysis, participants being expected to agree on statements with 'consecutive weights' (*ibid.*: 118). The resulting scoring method was

cumbersome, and alternative methods were developed, using only positive agreement for counting. This new form was tested for reliability, at different times and with different groups, and for validity, by obtaining 'ratings of personal adjustment' of a sample of participants by peers, by a set of 'judges', and by self-report, and correlated with the scores obtained by the inventory itself. This inventory, along with one focusing on 'external measure [of] the degree to which an individual is able to participate in the activities typical of adults' (ibid.: Examiner's Manual, 1), was filled in by approximately 5000 participants, manually processed and intensively analysed. The whole process of scale development and validation, data collection, and analysis took six years.

Usually taken to be a key study in the establishment of the concept of and policies promoting 'successful ageing' (Havighurst, 1963; Katz, 2000; see also Achenbaum, 1995: 106), PAOA articulated a view of transition to post-work life underpinned in large part by the capacity to mentally change one's self-conception and expectations, in adaptation to new social roles. Factors that facilitated that *successful adaptation* were, according to the PAOA team, anticipation, preparation, and control of the process (Burgess *et al.*, 1949: 27–8). The process of transition to retirement was one where there was

no single driver of satisfaction with old age (income, gender, marital status, health), but [the data showed] that these could be managed through psychological processes involved in the ageing person's adaptation of his [*sic*] attitudes. (ibid.: 74)

The social and economic landscape that underpinned the PAOA study, however, had changed in the years it had taken to develop the Cavan Inventory. Since 1935, US states had been incentivised to provide their own retirement programmes, supported by match funding by the federal government until around 1950, when the establishment of the tax-funded federal programme of Old Age Insurance led to an expansion of the system. On the other hand, regulation of wages attempting to contain wartime inflation since 1942 had increasingly led firms to secure labour by offering pension schemes, leading to a sixfold increase in the number of people with occupational retirement pensions between 1940 and 1960 (Costa, 1998). The growth in retirement at the turn of the 1950s was thus a central social and economic uncertainty, there being little data on how this would impact the well-being of retirees.

It was to address this uncertainty about the effects of the rise of pension schemes that Milton Barron drafted a research proposal on 'the impact of occupational retirement in the US on Physical and Mental morbidity and mortality'.² Its drafting was thus done in parallel with faculty officials seeking interest from funders, such as the National Institute of Mental Health or a private foundation, on the topic. Then an assistant professor at the Department of Sociology and Anthropology at Cornell, Barron had up to this point been focused on studying the position of ethnic and religious minorities in the US, drawing on the tradition of the 'Chicago School'. At the end of the 1940s, Barron had become interested in researching and teaching 'juvenile delinquency'. In this context, his 'impact of retirement' proposal was not a straightforward translation of his previous research interests. Its origin, although not documented, can partly be seen as linked to Lilly Endowment Fund's shift towards becoming a more active, professional funder

under the direction of Nick Hoyes, a Cornell alumnus and close associate of Elly Lilly himself.

Barron had framed his proposal by hypothesising that the mental and physical morbidity following retirement resulted from a lack of normative institutional support, retirement not serving the central American value of economic efficiency (see above; also Cowgill, 1974; Frank, 1946). In a manner similar to Cavan's formulation of the drivers of suicide (see above), Barron suggested that morbidity was a process of social disorganisation, leading to individuals' inadequacy, confusion, and suffering as an embodiment of 'society's tensions and cultural inconsistencies' (Barron, Streib, and Suchman, 1952: 479). He proposed to focus on occupational retirement schemes, studying participants' 'moral status' and personality before and after retirement through a baseline and follow-up study. In this respect, Barron's proposal was similar to the design of other – subsequently called longitudinal – studies of ageing at the turn of the 1950s, although its methodological justification makes no reference to these.³

Hoyes was receptive to the proposed focus on the effects of retirement on self and personality but suggested that Barron should also focus on the 'effects on the public economy'. This was because, as he wrote to Asa Knowles – then Cornell's vice president – 'control of [pension schemes] by the unions plus liberalization of old age allowances ... encourages extravagances and a lack of saving by working people during their active years'.⁴ Barron and colleagues were not convinced by this pre-emptive interpretation of their proposed study, however. As a compromise, Cornell and Lilly Endowment agreed to deepen the study's focus on health and health care, as this was a key component of the 'burden' Hoyes identified in occupational retirement schemes.

These negotiations also entailed enrolling two other faculty members at the Department of Sociology with methodological expertise. Gordon Streib, still an instructor, had previously been a member of the Bureau of Applied Social Research at Columbia University, directed by Paul Lazarsfeld, working under the guidance of Leo Löwenthal on audience research. As a fellow of the Social Science Research Council, he had developed innovative questionnaire methodologies with Navajo groups. Suchman, a senior colleague of Streib, had been trained in experimental psychology and psychometrics at Cornell's Psychological Laboratory in the late 1930s, and had conducted radio audience research, also under Lazarsfeld, before joining the Armed Forces Research Branch, where he worked most closely with Guttman (see above). An assistant professor at the time, Suchman was brought into the CSOR institutionally as a broker and academically as a 'methodological consultant' due to his role in leading the Social Science Research Center, itself in a key stage of development, having received a major five-year grant from the Ford Foundation (Cornell University, 1955).

The development of the instruments to be used in the study was thus a priority, and in early 1951 Suchman and Barron visited key research centres in the domain of public health and industrial relations.⁵ At Columbia University Teachers' College, Irving Lorge, while 'evasive', provided them with a copy of the schedule he was using in a study of adjustment to retirement with Jacob Tuckman, which included items on 'reported health status' (Tuckman and Lorge, 1953). Theodore Woolsey and Selwyn Collins, at the National Center for Health Statistics, questioned the premises of the study (see Woolsey, 1952), and suggested the CSOR team 'construct [their] own index

of health, which would involve a check list of illness and complaints'.⁶ In the final proposal to Lilly Endowment, the CSOR team stated their intention to combine the Cornell Medical Index (1949), to which Lorge had contributed, to be administered by physicians, and a selection of self-reported questions on satisfaction, religiosity, social relations, working conditions, plans, and health, these latter combining the Tuckman-Lorge schedule with some items from the Cavan Inventory (see above).

With these guarantees, Lilly Endowment agreed to provide \$130,000 to support a seven-year study, and the long process of recruiting companies and assembling the sample of participants commenced. This also involved expanding the team to include a person responsible for analysing the Cornell Medical Index data, 'ideally equipped to "translate" the medical findings in terms meaningful to social science' for the 'in-plant' studies of retirement (see below).⁷ The aim was for this data to provide professionally certified evidence of morbidity, but also to address the problem of validity in the same way that, in the Research Branch studies, 'the army handled the problem of evaluating the psychiatric status of the men by having independent analyses done by psychiatrists and screening questionnaires'.⁸ In this, the measurement of adjustment at the CSOR was focused not only on satisfaction and happiness, as the PAOA team had done, but also sought to provide evidence of somatic adaptation to retirement. Further, it suggested that 'attitudes and activities' could hardly ever 'replace gainful employment for a retirant [*sic*]' (Barron, Streib, and Suchman, 1952: 481; see also Barron, 1952), and that socio-economic status (SES) and social networks could play an important role in adjustment processes.

This shift towards a more 'objectivist' orientation in the study was accompanied by changes in its management, with Streib becoming a co-director in 1953, Suchman consolidating his central role as a 'specialist with regard to the methodological problems of the study', and Wayne Thompson, a student of Streib, becoming its main researcher ('Field Director'). This also marked the shift to a more streamlined organisation of the recruitment and data collection operation, such that by end of 1953, the study could count on the participation of 340 organisations.⁹ In addition, data entry and processing, using IBM punched-card machines, started facilitating the writing of preliminary reports, providing the team with some confidence in the quality of the data in both the survey component of the study and the 'in-plant' studies. In early 1954, Barron left Cornell and the CSOR, and Streib became director of the study, driving its data operation.

Much of 1954–5 was focused on data collection, entry, and processing, and analysis of the baseline data. A major component of this was scale analysis, using sorting programs specially designed for the IBM machines at the Cornell computing laboratory, to emulate the scalogram developed by Guttman and Suchman at the Armed Forces Research Branch (see above). For example, for the teams' participation in the 3rd International Gerontological Congress in London, they developed their own approach to the measurement of adjustment based on the scale pattern formed by questions on goal-centeredness, satisfaction, and reaction to adversity. These items 'combined and ordered according to the Guttman scale model [obtained] a coefficient of scalability [reproducibility] of .95' (Streib, 1956: 272).¹⁰

Crucially, this represented a different approach to measuring adjustment to the one developed in the PAOA study. Contrary to the PAOA team, who had mostly disregarded SES in their sampling (Burgess *et al.*, 1949: 50–4), the CSOR team had been able to

include employees across the salary scale in the companies recruited, supporting their validation of the self-reported SES scales and their measurement of the (positive) statistical relationship between SES and 'morale in the retired'. With this, the CSOR team challenged the assumptions of adjustment to retirement research, showing that the drivers of satisfaction were primarily income and 'health' (see below). As Thompson expressed in his doctoral dissertation, 'Evidently, retirement is not as disruptive of the personality as it has been frequently thought to be', provided retirees have access to material/somatic resources (Thompson, 1956: 132).

This focus on what Thompson and Streib later labelled 'situational resources' would come to define the CSOR approach to the analysis of transition to retirement (e.g. Thompson and Streib, 1958). The CSOR team pointed to the *patterened* configurations underpinning 'role flexibility', and the differential ability to take advantage of the opportunities of retirement. In a direct challenge to the Havighurst model, Thompson and Streib (1958) questioned whether preparation and planning for retirement were, by themselves, key in bringing about successful adjustment (see also Thompson, Streib, and Kosa, 1960). It was not retirement itself or its various forms that represented a problem, but the economic circumstances, health, and 'social situation' in which the decision to retire came about that determined successful transition to 'old age'. This conclusion also diverged from the expectation the study's main funder had had for the analysis (see above).¹¹

As expected, both reacted, with the study's main funding being transferred to the NIH, and Havighurst publicly criticising the CSOR at the 1959 Gerontological Society of America (GSA) Conference at Ann Arbor for seeking 'to examine only medical and economic problems, therefore [being] beyond [its] data to examine adjustment'.¹² Responding to a request for clarification by Streib, Havighurst argued his criticism was directed at those who had used the Thompson and Streib (1958) paper to support doubt in the 'efficacy of retirement planning programmes', of which he had been a key proponent, and to suggest that higher incomes might be more important for good adjustment.¹³ His disagreement with the CSOR, he argued, was 'to do with the method used in the study to get information about planning for retirement', that is, its measurement procedures. Thompson, responding on behalf of the CSOR, explained that the item on planning was not attempting to measure adjustment but was 'a measure of specific plans made for retirement' (a behaviour) which was in a scalar relationship (in a configuration) with other behavioural predictors of adjustment.¹⁴

Their disagreement was thus about the *politics of method*. Havighurst, supported by prominent figures in gerontology such as James Birren, conceived of adjustment as a 'psychological continuum', on which the position of each individual could be determined by a series of questions-items with varied weights.¹⁵ Adjustive behaviours, of which planning for retirement was one (see above), could be measured with greater certainty if based on a large set of items. For the CSOR team, planning for retirement was a condition in a configuration of behaviours to be taken in their entirety, per Guttman's 1941 proposal (see above). The question of whether it was an adjustive behaviour was empirical, to be answered by scale analysis. As it turned out, this item was in a scalar relationship with positive attitude to retirement, making it part of an arrangement of conditions that enhanced retirement as an opportunity for action. These presented different loci of

intervention, Havighurst aiming to shift the individual along the ‘psychological continuum’ through planning for retirement programmes, and the CSOR supporting a more nuanced focus on the resources available for the individual to act on. This disagreement about the ontological politics of measurement was even more striking in relation to health, as we will see in the next section.

Struggling with the meaning of health

From its inception, the CSOR deployed a strain between different components of the work of the Cornell Sociology Department. On the one hand, the proposal fitted well within Knowles’ and the wider university management’s aim to obtain more external funding for the social sciences after the war. By focusing on retirement and morbidity/health, Barron was able to capture Lilly Endowment Fund’s interest, as was described in the last section. This represented a strategic alignment with an ongoing public debate about health care, as just two years before President Truman had proposed to create a system of universal health insurance coverage. Such proposals, as documented by Oberlander (2003), were met with opposition from the American Medical Association, which, in coalition with the Republican Party, was publicly critical of ‘socialised medicine’, as it undermined the value of ‘choice’ and freedom.

Barron’s original proposal drew on a vague hypothesis that retirement caused mental disorders due to ‘social disorganisation’ through loss of status, habits, and institutional support (see above), and contained no details of the instruments to be used in the ‘Nationwide Survey’ or the ‘Follow-up, Situational Studies’ it outlined. In relation to morbidity and health, drawing heavily on then ongoing research in the Social Science Research Center, he added:

It is not within the scope of this study to examine in detail the diseases of old age. But the question of what is the relationship between declining health and adjustment to old age is a crucial one?... What is the effect of ... frequent chronic ailments in adjustment to old age? Are those who have a history of illness more likely to make a satisfactory adjustment?... To what extent is such adjustment a matter of a person’s knowledge, of his general health habits, of the expectations of the social stratum, of his occupational role?¹⁶

As we know, one of the first tasks undertaken by Suchman as a ‘methodological consultant’ in aiming to secure the grant agreement was to conduct a tour of a variety of organisations to gather possible instruments to investigate these questions. Of these organisations, none were social science research centres: an actuary in an insurance company; Lorge, at Teacher’s College of Columbia University (see above); Woolsey and Collins at the US Public Health Service; the Bureau of Labor Statistics; the National Research Council; the Social Security Administration; the World Health Organization; and the US Census Office.¹⁷ Reviewing the schedules and instruments supplied to the CSOR team, Suchman was still uncertain about their expertise in handling and interpreting the health data, and suggested, based on his experience at the Research Branch, that they involve a public health expert to deal with the medical examination data and the health components of the ‘sociological survey’.¹⁸

The uncertainty experienced by the CSOR team in relation to health was not unique. In the PAOA study, health had been introduced in the inventory latterly as a result of the analysis of the interviews with informants in the first phase of the development of statements for the questionnaire (see above). Although it is the first recorded question-answer form of what later became known as 'self-rated health' (see introduction), in the Cavan Inventory, the general health question is defined as a measure of 'attitude towards health' (Burgess *et al.*, 1949: 56) – or an 'affective reaction to the situation' (*ibid.*: 91) – but was integrated into the activities section of the schedule. This insertion, however, was also inaccurate from the point of view of the PAOA team, who added that 'the health questions are not, properly speaking, questions about activities, but since health is closely related to many of the activities of older people, the health questions were included [there]' (*ibid.*: 137). Further, and significantly, the health items were scaled using a Likert-like technique, rather than the Thurstonian 'agree/disagree' system that was used for the rest of the schedule.

Drawing on his experience at the Armed Forces Research Branch in comparing scaled questionnaire items with psychiatric assessment (see above), and Kleemeier's (1951) experience of validating the health questions of the Cavan Inventory in a residential facility for older people with physicians' estimates, Suchman proposed a collaboration with the Department of Public Health at Cornell. This led him to approach Emmerson Day, a specialist in preventative medicine with whom he had collaborated on a study of cancer diagnosis. Day, however, was himself ambivalent about this collaboration, despite Suchman's impressing on him that, while the medical component of the study was 'previously a subsidiary element to the study, [it had] become one of its major aspects'.¹⁹ Day's reluctance to commit made the CSOR 'an impossible research operation',²⁰ because without his involvement, the team would have to acknowledge 'the limitations of the questionnaire technique' in assessing health.²¹

When, eventually, Day agreed to process and analyse the medical forms and Cornell Medical Index data collected in the 'in-plant' study of the CSOR, Streib and Suchman were not satisfied with the level of statistical proficiency of the analysis: it included no percentages or tables, for example.²² It was thus impossible to compare, even roughly, the medical data with the questionnaire data. In the first report on the project, Barron and colleagues presented a tentative, unvalidated analysis of participants' self-appraisal of their health, comparing employees in the different types of industry involved in the study.²³ A further analysis, performed by Streib himself based on Day's numerical data on 'the completed medical records', revealed

an interesting finding [in] the doctor's rating of the subject's health... It is interesting to note that percentage wise approximately the same proportion of people fall in the various categories as was found in a preliminary run of some 850 cases in which the subjects rated their own health. We would need to cross tabulate the individual cases according to their own evaluation and the doctor's rating in order to ascertain the accuracy of a person's subjective evaluation of his health. However, the overall distribution indicates a pretty fair correspondence.²⁴

This provided the CSOR team with some confidence that the health items included in the questionnaire related to significant phenomena, and a paper relating 'occupational roles'

to health was outlined and planned to be presented at the American Sociological Association (ASA) conference that year. Barron's withdrawal in 1954 did slow the team's engagement with the health data, but enabled the enrolment of Bernard Kutner, a social psychologist from the Public Health Department at Cornell who worked with the Social Science Research Center. This meant the CSOR team had found their 'translator' of medical data, increasing their confidence that the scale patterns they were seeing in the questionnaire data had some 'correlate' with physicians' assessment. Underpinned by this, Suchman undertook a major scalogram analysis of the items used in the survey at the end of 1954, identifying a scalar pattern in the health questions with a high coefficient of reproducibility (see above), leading to the conclusion that 'self-administered questionnaires are useful in obtaining health and medical information'.²⁵ In this assessment, Guttman techniques were of key importance:

We find that these three items [Has your health changed during the past year? / How would you rate your health at the present time? / Do you have any particular physical or health problems?] fit together into a scalar relationship, i.e. a given answer on one of them enables us to predict with reasonable confidence the responses to others, thus providing a validity check on the content of each and making it possible to distinguish with greater accuracy ... people [in relation to their] self-health evaluation.²⁶

The CSOR team had, in a way, put a fence around a mystery. They knew they could measure/scale health items, and that, as a consequence, self-assessed health had 'meaning' in the methodological terms defined by Suchman (1950) himself, but remained unsure what the meaning was in sociological terms: was health an attitude, a belief, or something else? This uncertainty was, however, different from that experienced by the PAOA study team, because the scale used by Cavan and colleagues was a simple operationalisation, transforming health into a 'psychological continuum' to enable measurement. For the CSOR team, being able to scale health meant it obtained as an object. In this, the self-evaluation of health was a configuration of 'a class of behaviours' and not merely a single item in a schedule, as it was in the Cavan Inventory. To fully form the object, however, they needed to be able to emplace it more firmly in a coherent cognitive and political schema (Desrosières, 1990).

This was a challenge, as there was no stable, identifiable set of concepts and institutions to refer to: medical sociology was then only emerging as an area of research and teaching in the US, but not at Cornell. This was reflected in the CSOR team's decision to commit to a different area of research. Thus, proposing 'a comprehensive program of research on aging and retirement' for funding by Lilly Endowment in early 1955, Streib suggested focusing on 'sociological areas of interest' and reducing the emphasis on health and 'diseases of old age'.²⁷ Further, when presenting the study's health survey data to the ASA in the same year, Streib claimed that deploying health as a variable had 'ample theoretical justification', which he did not provide, however, apart from referring the audience to Parsons' *Social System* (Streib, 1956: 271).

Some conceptual development came in the form of the idea of 'situational determinants', in which health could be thought of as both a resource and an 'orientation' or evaluation that the actor makes of those resources (Thompson, 1956; Thompson and

Streib, 1958). Thus, it was possible to measure health as a ‘situational resource’ – as *objective health* in the form of physicians’ ratings, and as *subjective health* as ‘evaluated by the individual himself [*sic*]’ (Streib and Thompson, 1957) – and to relate them through the conceptual device of a strategic subject. This enabled the CSOR team to claim that an individual’s position on the health scale predicted, with some certainty, his/her decision to retire, and was not a consequence of it, which underpinned their position on the retirement adjustment controversy (see above). It did not, however, drive the team immediately to engage with the emerging field of medical sociology, because from their perspective these ‘findings [were still] subject to the charge that individuals’ self-evaluation of health is not “really” an adequate index’ of actual health.²⁸

Again, the team’s underlying uncertainty about the cognitive/epistemic meaning of their ‘subjective health scale’ and how it related to physiological conditions continued, despite it being the scale with the highest reproducibility coefficient in the entire data set (e.g. for the 1954 data, CR = .96). This uncertainty was also closely related to the diversity of ways in which subjective health could be related to other variates. Was it a ‘proxy’ measure for actual health, or a dependent variable for changes in occupational status, or an intervening variable in retirement decisions, or a predictor of attitudes? There seems to have been some disagreement within the team about which way the scale should be deployed, Thompson preferring the former approaches, and Suchman focusing increasingly on its predictive status for action, subjective health being ‘a better predictor of health behaviour such as staying at home from work because of illness or visiting the doctor *than clinicians’ reports*’.²⁹ This predictive value was, however, not enough to persuade his sociologist peers, and Suchman struggled to get his paper on the subjective health scale accepted in mainstream sociological journals, the *American Sociological Review*’s ‘editor [for example] suggest[ing] that it might be more appropriate for a public health publication’.³⁰

By the time of the publication of the paper usually cited as the first publication on SRH (Suchman, Phillips, and Streib, 1958), Suchman had become Director of Social Science at the New York City Department of Health, and Streib could confirm to the NIH that the study would abandon its health focus entirely to emphasise ‘sociological variables’.³¹ This involved tapering off the ‘medical programme’, the data of which remained largely unanalysed, and the health components of the follow-up questionnaires. From his subsequent position at the University of Pittsburgh, where he went on to develop a research programme in medical sociology (Suchman, 1963), Suchman considered that both the subjective health scale and the associated data belonged to Streib as director of the CSOR, and did not use this data again in his work. Thompson himself planned to move on to ‘new work in political sociology’.³² Streib went on to focus on retirement and the family, abandoning the focus on health completely.

Detaching from the health aspects of the study was not a straightforward process, however. Following the claim by the American Medical Association (AMA) that a study by Wiggins and Schoeck (1961) demonstrated that the US medical system provided sufficient care for the elderly, and thus weakened claims by unions and Democrats that a health service for older people might be necessary, Thompson and Streib were asked to comment by the International Association of Gerontological Societies Conference in 1960. The Wiggins study had been of use in the political controversy because it

claimed, based on SRH data, that Americans tended to relatively underrate their health status, possibly leading to unnecessary health care usage. Thompson, however, claimed that the CSOR SRH data showed instead that “‘fair health’ fits with “‘poor’ and “‘very poor’ as validated by the Guttman scaling technique, and correlated with physical examinations and predictive tests of validation’.³³ Rather than undervaluing their health state, ‘as many as two out of three who were rated “unfavourable” by the physician gave themselves “favourable” reports’ (Suchman, Phillips, and Streib, 1958: 226), suggesting a process of accommodation in view of the lack of accessible health services.

Defending their method against those who, like Senator McCarthy (Dem, Minnesota), had reacted to the AMA’s use of the Wiggins and Schoeck study by questioning ‘the ability of an individual in an interview with sociologists to determine the actual state of his physical or mental condition’,³⁴ Thompson suggested that ‘subjective health’ data instead indicated reluctance to use medical services among older Americans, because of how its components scaled as a ‘configuration’ of a ‘class of behaviours’. Here, again, the *politics of method* was important. Using the single psychometric SRH scale that Cavan and colleagues had used, Wiggins and Schoeck had considered ‘fair health’ to be an inherently positive/favourable rating, but the CSOR Guttman scalogram analysis had clearly positioned it as closely related to other unfavourable items (‘poor’ and ‘very poor’). This enabled emplacing subjective health in a political frame of reference and articulating it as a measure of trust in older Americans’ capacity to reliably perform help-seeking behaviour in a federally supported health care system. SRH had found its temporary political framing as a predictor of ‘health seeking behaviour’, and a driver of health care demand.

Conclusion

In this article, I have proposed a genealogy of the scalable subject, focusing on health and, in particular, the case of SRH, one of the most widely used metrics for management of populations and individuals in contemporary societies. This article’s point of departure was the suggestion that the development of SRH was a window to understanding how measurement came to shape and be shaped by the institutions that deploy it through a series of ‘comparisons, negotiations, compromises, translations, inscriptions, codings, of codified and replicable procedures and calculations’ (Desrosières, 2008). The article detailed these negotiations and translations through three nested layers of uncertainty and argued that SRH came into existence at the *contingent but generative*, unlikely intersection of those different planes.

The first uncertainty concerned scaling methodology in psychological and social research. The article traced how the ‘metaphysical’ assumptions that enabled psychologists like Thurstone to develop techniques to quantify individuals’ qualities were exposed and challenged by Guttman’s approach to data and prediction, drawing heavily on the computational resources and practical requirements of the Armed Forces Research Branch. In particular, this section suggested that Guttman’s proposal was attractive to social scientists not only because ‘it dispense[d] with the concept of [an] underlying [psychological] continuum’ (Stouffer, 1950: v), but also because it presented an elegant, practical solution to the emerging ‘social engineering’ mission of the

social sciences, where the cumbersome problems of item selection and weighting would be 'actually non-existent' (Suchman, 1950: 80).

The second section focused on the activities of the CSOR and its interaction with ongoing academic and public debates about the effects of retirement on personality, health, and social organisation. Identified as a challenge to modern American 'pragmatic culture', retirement had, since the late 1930s, been seen by social scientists as an 'experiment' in social adjustment requiring both normative, institutional change and individual attitudinal adaptation. This section detailed how, through the compromises between the goals of the academics and those of their sponsors, the CSOR came to combine the Bureau of Social Research and Armed Forces Research Branch forms of reasoning and calculation practices and adapt them to the public issue of retirement. Addressing both academic and political debates, the CSOR models suggested that health and income regulated individuals' decisions to retire, challenging central epistemic, methodological, and political assumptions about understanding and managing adjustment to post-work life.

It is from these two controversies that it is possible to understand how SRH came into existence. Initially intending to rely on comprehensive medical examination data, the CSOR team had to make sense of their own health data without much recourse to clinical expertise. Drawing on Suchman's experience in developing the computational procedures and visual techniques of scalogram analysis, the CSOR team were able to reveal a strong 'scalar relationship' in the health items included in the participants' self-administered questionnaire. This was crucial in turning SRH into an object, a 'measure of something' not yet fully defined, but provisionally called subjective or 'perceived health' (Suchman, Phillips, and Streib, 1958: 232). Being reluctant to invest in health as a domain of research, the CSOR team were gently incited to focus on their health items because of their 'Guttman scalability' and how well they correlated with a variety of dependent variables. However, even with such methodological assurance, the team struggled to agree on its sociological meaning and to enrol editors and peer reviewers, with their study finding its place in public life only as a reference within debates about the political organisation of health care.

In short, I am proposing that Guttman scalogram analysis was integral to making health into a measurable, tractable object. The version of the history of SRH that I am submitting differs considerably from the established account, which emphasises the role of psychometrics in turning health into a quantifiable subjective quality. As I have suggested, while the first recorded use of the SRH question-answer format was in the Cavan Inventory, the PAOA team could not place health within any 'coherent cognitive or political schemas' – either attitudes or activities – and this meant that they were effectively unable to measure health in any meaningful way. Maddox's (1962) interpretation of the health items in the Cavan Inventory as 'self-assessment of health' in the Duke Longitudinal Study of Aging does not reference the CSOR 1958 paper, instead making use of the concept of 'health self-rating' developed by Bernard Kutner, who himself had drawn on his own collaboration with the CSOR team to generate that notion (see above; also Kutner *et al.*, 1956). This is not surprising given Maddox's adherence to Havighurst's model of adjustment to old age.

Indeed, I am also suggesting that the disputes and tensions – the *politics of method* – that led to the emergence of SRH did not get resolved by its establishment. Using

Latour's conceptual terminology, it could be said that SRH became a mediator – rather than an intermediary – in the making of the scalable subject, 'transforming, translating, distorting, and modifying the meaning' it was supposed to carry (Latour, 2005: 39). Thus, as the second and third sections of this article made clear, SRH became particularly relevant in the academic-public debates about the extension of health insurance to older populations in the US. In this, there was also debate and reinforcement of the methodological expertise of social scientists 'in obtaining health and medical information' from questionnaires, a question that became more significant as medical outcome measurement and health service quality indicators became mainstream from the 1980s onwards (Elwood, 1988). While a standard metric in health care systems, population surveys, and digital health apps, uncertainty remains central to the deployment of SRH, and thus research on the bio-psycho-social mechanisms that might explain – translate, or transform – SRH continues to the present day.


Declaration of conflicting interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Tiago Moreira  <https://orcid.org/0000-0002-6950-865X>

Notes

The rough outline of this article was first presented at the symposium 'Experience as Evidence', University of Oxford, 13–14 October 2014. I thank Fahila Mazandenari for inviting me to speak at the workshop, and the other participants for their comments and suggestions. Versions of the article were also presented at the Centre for Medical Science and Technology Studies (June 2015), the London BSA Medical Sociology Group, LSHTM (June 2016), the Cambridge CRASSH Workshop (July 2016), and Tomsk State University (June 2017). Archival data collection was supported by Durham University, and generously hosted by the Inter-university Consortium for Political and Social Research, University of Michigan, and by Amy Pienta. The article has benefited from generous comments from Jana Bacevic, Stephen Casper, Stephen Katz, Henriette Langstrup, and Lucy Suchman.

1. The CSOR data is archived at the Inter-university Consortium for Political and Social Research. References to this resource are coded as [CSOR, Box #: Author, title, place of publication, date, page number]. Many thanks to Amy Pienta for her generosity in facilitating and advising me on access to the archive.
2. CSOR, Box 3: Barron, 'The Impact of Retirement', February 1950.
3. CSOR, Box 13: Barron, 'Proposal for the Establishment of an Institute of Gerontology and Geriatrics at Cornell University', February (?) 1952.
4. CSOR, Box 12: Hoyes to Knowles, 25 October 1950.

5. CSOR, Box 12: 'Report on Trip to New York, Newark and Washington', 28 January–3 February 1951.
6. 'Report on Trip to New York, Newark and Washington', p. 3.
7. CSOR, Box 14: Suchman, 'Proposal for the Appointment of Medical Research Fellow to CSOR', 1952, p. 20.
8. CSOR, Box 12: Suchman (?), 'Luncheon Meeting Memo', 25 July 1951.
9. CSOR, Box 15: Streib to Dulling, at Lilly, 6 October 1953.
10. CSOR, Box 15: Suchman, 'The Reliability of Scales: An Analysis of Scales Employed in the Retirement Study', 1955.
11. CSOR, Box 16: Streib to Patillo, 11 February 1958.
12. CSOR, Box 16: Thompson to Streib, 25 November 1959.
13. CSOR, Box 21: Havighurst, 'Memorandum Concerning the Cornell Retirement Study', 18 December 1959.
14. CSOR, Box 21: Thompson to Havighurst, December 1959.
15. CSOR, Box 16: Thompson to Streib, 25 November 1959.
16. Barron, 'The Impact of Retirement'.
17. 'Report on trip to New York, Newark and Washington'.
18. CSOR, Box 12: Suchman (?), 'Luncheon Meeting Memo', 25 July 1951.
19. CSOR, Box 16: Suchman to Day, 28 May 1952.
20. CSOR, Box 16: Suchman to Day, 7 July 1952.
21. CSOR, Box 16: Irving Ladimer, NIH, to Barron, 29 September 1952.
22. CSOR, Box 16: Streib to Day, 10 July 1953.
23. CSOR, Box 12: Barron, 'The Study of Occupation Retirement', Social Science Research Center, Ithaca, 1953.
24. CSOR, Box 21: Streib to Day, 19 February 1954.
25. CSOR, Box 15: Suchman, 'Summaries of Major Findings to Date', March 1955, p. 6.
26. CSOR, Box 21: Thompson to Harry Johnson, at Life Extension Examiners, 13 March 1957.
27. CSOR, Box 16: Streib, 'A Proposal for the Developments of a Comprehensive Program of Research on Aging and Retirement', February 1955.
28. CSOR, Box 21: Thompson and Streib, 'Health and Retirement', manuscript of paper presented at GSA, Cleveland, 1957.
29. CSOR, Box 21: Suchman in Streib, 'Progress Report Submitted to Division of Research Grants NIH', April 1957, p. 3 (emphasis added).
30. CSOR, Box 22: Streib to Phillips, 22 April 1957.
31. CSOR, Box 22, Streib, 'Progress Report Submitted to NIH Division of Research Grants', July (?) 1958.
32. CSOR, Box 23: Streib to Thompson, 10 November 1959.
33. CSOR, Box 23: Thompson, 'Comment in 5th Congress of the IAG', August 1960.
34. McCarthy, Congressional Record Daily Digest, 22 August 1960: CD453.

References

- Abbot, A. (1999) *Department and Discipline: Chicago Sociology at One Hundred*. Chicago, IL: University of Chicago Press.
- Achenbaum, W. A. (1995) *Crossing Frontiers: Gerontology Emerges as a Science*. Cambridge: Cambridge University Press.

- Arbel, T. (2016) 'The American Soldier in Jerusalem', PhD thesis, Harvard University.
- Barron, M. (1952) 'Involuntary Retirement and Morbidity and Mortality', *Public Health Reports* 67(2): 128.
- Barron, M., Streib, G., and Suchman, E. (1952) 'Research on the Social Disorganization of Retirement', *American Sociological Review* 17(4): 479–82.
- Bowling, A. (2005) 'Just One Question: If One Question Works, Why Ask Several?', *Journal of Epidemiology and Community Health* 59(5): 342–5.
- Bowling, A. (2017) *Measuring Health: A Review of Quality of Life Measurement Scales*. Milton Keynes: Open University Press.
- Bulmer, M. (1980) 'The Early Institutional Establishment of Social Science Research: The Local Community Research Committee at the University of Chicago, 1923–30', *Minerva* 18: 51–110.
- Burgess, E. W. (1916) 'The Social Survey: A Field for Constructive Service by Departments of Sociology', *American Journal of Sociology* 21(4): 492–500.
- Burgess, E. W. (1934) *The Adolescent in the Family: A Study of Personality Development in the Home Environment*. New York, NY: Appleton-Century-Croft.
- Burgess, E. W., Havighurst, R. J., Cavan, R. S., and Goldhamer, H. (1949) *Personal Adjustment in Old Age*. Chicago, IL: Science Research Associates.
- Carson, J. (2007) *The Measure of Merit: Talents, Intelligence, and Inequality in the French and American Republics, 1750–1940*. Princeton, NJ: Princeton University Press.
- Cavan, R. S. (1928) *Suicide*. Chicago, IL: University of Chicago Press.
- Converse, J. M. (1987) *Survey Research in the United States: Roots and Emergence*. Berkeley, CA: University of California Press.
- Cornell University (1955) *The 1950 Ford Grant to Cornell: A Report and an Evaluation*. Ithaca, NY: Cornell University, Social Science Research Center.
- Costa, D. L. (1998) *The Evolution of Retirement: An American Economic History, 1880–1990*. Chicago, IL: University of Chicago Press.
- Cowgill, D. O. (1974) 'The Aging of Populations and Societies', *Annals of the American Academy of Political and Social Science* 415(1): 1–18.
- Danziger, K. (1990) *Constructing the Subject: Historical Origins of Psychological Research*. New York, NY: Cambridge University Press.
- Danziger, K. (1997) *Naming the Mind: How Psychology Found Its Language*. Thousand Oaks, CA: SAGE.
- Desrosières, A. (1990) 'How to Make Things Which Hold Together', in P. Wagner, B. Wittrock, and R. Whitley (eds) *Discourses on Society: The Shaping of the Social Science Disciplines*. Dordrecht: Springer, pp. 195–218.
- Desrosières, A. (2008) *Pour une sociologie historique de la quantification* [Towards a Historical Sociology of Quantification]. Paris: Presses des Mines.
- Elwood, P. M. (1988) 'Outcomes Management: A Technology of Patient Experience', *New England Journal of Medicine* 318: 1549–56.
- Falconer, J. and Quesnel-Vallée, A. (2017) 'Pathway From Poor Self-Rated Health to Mortality: Explanatory Power of Disease Diagnosis', *Social Science & Medicine* 190: 227–36.
- Foucault, M. (1991) *The Foucault Effect: Studies in Governmentality*, ed. G. Burchell, C. Gordon, and P. Miller. Chicago, IL: University of Chicago Press.
- Frank, L. K. (1939) 'Foreword', in E. V. Cowdry (ed.) *Problems of Ageing*. Baltimore, MD: Williams & Wilkins, pp. i–v.

- Frank, L. K. (1946) 'Gerontology', *Journal of Gerontology* 1: 1–12.
- Guttman, L. (1941) 'The Quantification of a Class of Attributes: A Theory and Method of Scale Construction', in P. Horst, P. Wallin, L. C. Guttman, F. B. Wallin, J. A. Clausen, R. Reed, and E. Rosenthal (eds) *The Prediction of Personal Adjustment*. Chicago, IL: Social Science Research Council, pp. 321–48.
- Guttman, L. (1942) 'A Revision of Chapin's Social Status Scale', *American Sociological Review* 7(3): 362–9.
- Guttman, L. (1944) 'A Basis for Scaling Qualitative Data', *American Sociological Review* 9(2): 139–50.
- Guttman, L. (1947) 'The Cornell Technique for Scale and Intensity Analysis', *Educational and Psychological Measurement* 7(2): 247–79.
- Guttman, L. and Suchman, E. A. (1947) 'Intensity and a Zero Point for Attitude Analysis', *American Sociological Review* 12(1): 57–67.
- Havighurst, R. J. (1963) 'Successful Aging', *Processes of Aging: Social and Psychological Perspectives* 1: 299–320.
- Havighurst, R. J., Kuhlen, R. G., and McGuire, C. (1947) 'Personality Development', *Review of Educational Research* 17(5): 333–44.
- Hoeyer, K., Bauer, S., and Pickersgill, M. (2019) 'Datafication and Accountability in Public Health: Introduction to a Special Issue', *Social Studies of Science* 49(4): 459–75.
- Idler, E. L. and Benyamini, Y. (1997) 'Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies', *Journal of Health and Social Behavior* 38: 21–37.
- Igo, S. E. (2007) *The Averaged American: Surveys, Citizens, and the Making of a Mass Public*. Cambridge, MA: Harvard University Press.
- Jylhä, M. (2009) 'What Is Self-Rated Health and Why Does It Predict Mortality?', *Social Science & Medicine* 69(3): 307–16.
- Katz, S. (1996) *Disciplining Old Age: The Formation of Gerontological Knowledge*. Charlottesville, VA: University Press of Virginia.
- Katz, S. (2000) 'Busy Bodies: Activity, Aging, and the Management of Everyday Life', *Journal of Aging Studies* 14(2): 135–52.
- Kleemeier, R. W. (1951) 'The Effect of a Work Program on Adjustment Attitudes in an Aged Population', *Journal of Gerontology* 6(4): 372–9.
- Kutner, B., Fanshel, D., Togo, A. M., and Langner, T. S. (1956) *Five Hundred Over Sixty: A Community Survey on Aging*. New York, NY: Russell Sage Foundation.
- Latour, B. (1990) 'Drawing Things Together', in M. Lynch and S. Woolgar (eds) *Representation in Scientific Practice*. Cambridge, MA: MIT Press, pp. 19–68.
- Latour, B. (2005) *Reassembling the Social: An Introduction to Actor-Network-Theory*. New York, NY: Oxford University Press.
- Lupton, D. (2018) *Digital Health: Critical and Cross-disciplinary Perspectives*. London: Routledge.
- Lutz, C. (1997) 'Epistemology of the Bunker', in J. Pfister and N. Schnog (eds) *Inventing the Psychological: Toward a Cultural History of Emotional Life in America*. New Haven, CT: Yale University Press, pp. 245–69.
- Maddox, G. L. (1962) 'Some Correlates of Differences in Self-Assessment of Health Status Among the Elderly', *Journal of Gerontology* 17: 180–5.
- Maddox, G. L. (1987) 'Aging Differently', *The Gerontologist* 27(5): 557–64.

- Maranell, G. M. (1974) *Scaling: A Sourcebook for Behavioral Scientists*. San Francisco, CA: Transaction.
- McDowell, I. (2006) *Measuring Health: A Guide to Rating Scales and Questionnaires* (3rd ed.). Oxford: Oxford University Press.
- Moreira, T. (2017) *Science, Technology and the Ageing Society*. London: Routledge.
- Oberlander, J. (2003) *The Political Life of Medicare*. Chicago, IL: University of Chicago Press.
- Office for National Statistics (2013, 30 January) 'General Health in England and Wales: 2011 and Comparison With 2001', <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandwellbeing/articles/generalhealthinenglandandwales/2013-01-30>.
- Park, H. W. (2008) 'Edmund Vincent Cowdry and the Making of Gerontology as a Multidisciplinary Scientific Field in the United States', *Journal of the History of Biology* 41(3): 529–72.
- Pienta, A. M. and Lyle, J. (2018) 'Retirement in the 1950s: Rebuilding a Longitudinal Research Database', *IASSIST Quarterly* 42(1): 12, available at: <https://doi.org/10.29173/iq19>.
- Pollack, O. (1948) *Social Adjustment and Old Age: A Research Planning Report*. New York, NY: Social Science Research Council.
- Riley, M. W., Riley, J. W. Jr., and Toby, J. (1954) *Sociological Studies in Scale Analysis: Applications, Theory, Procedures*. New Brunswick, NJ: Rutgers University Press.
- Ryan, J. W. (2010) 'What Were They Thinking? Samuel A. Stouffer and the American Soldier', PhD thesis, University of Kansas.
- Schilling, R. and Casper, S. T. (2015) 'Of Psychometric Means: Starke R. Hathaway and the Popularization of the Minnesota Multiphasic Personality Inventory', *Science in Context* 28(1): 77–98.
- Schooler, C. (1968) 'A Note of Extreme Caution on the Use of Guttman Scales', *American Journal of Sociology* 74(3): 296–301.
- Stark, L. (2018) 'Algorithmic Psychometrics and the Scalable Subject', *Social Studies of Science* 48(2): 204–31.
- Stouffer, S. A. (1941) 'How a Mathematician Can Help a Sociologist', *Sociometry* 4(1): 56–63.
- Stouffer, S. A. (1950) 'Foreword', in S. A. Stouffer, L. Guttman, E. A. Suchman, P. Lazarsfeld, S. A. Star, and J. A. Clausen, *Studies in Social Psychology in World War II: Vol. 4. Measurement and Prediction*. Princeton, NJ: Princeton University Press, pp. v–vii.
- Stouffer, S. A., Guttman, L., Suchman, E. A., Lazarsfeld, P., Star, S. A., and Clausen, J. A. (1950) *Studies in Social Psychology in World War II: Vol. 4. Measurement and Prediction*. Princeton, NJ: Princeton University Press.
- Stouffer, S. A., Suchman, E. A., DeVinney, L. C., Star, S., and Williams, R. M. Jr. (1949) *Studies in Social Psychology in World War II: Vol. 2. The American Soldier: Combat and Its Aftermath*. Princeton, NJ: Princeton University Press.
- Streib, G. F. (1956) 'Morale of the Retired', *Social Problems* 3(4): 270–6.
- Streib, G. F. and Thompson, W. E. (1957) 'Personal and Social Adjustment in Retirement', in W. Donahue and C. Tibbitis (eds) *The New Frontiers of Aging*. Ann Arbor, MI: University of Michigan Press, pp. 194–5.
- Suchman, E. A. (1950) 'The Logic of Scale Construction', *Educational and Psychological Measurement* 10(1): 79–93.
- Suchman, E. A. (1963) *Sociology and the Field of Public Health*. New York, NY: Russell Sage Foundation.

- Suchman, E. A. and Francis, R. (1954) 'Scaling Techniques in Social Research', in *An Introduction to Social Research*. Harrisburg, PA: Stackpole, pp. 126–9.
- Suchman, E. A., Phillips, B. S., and Streib, G. F. (1958) 'An Analysis of the Validity of Health Questionnaires', *Social Forces* 36: 223–32.
- Thompson, W. E. (1956) 'The Impact of Retirement', PhD thesis, Cornell University.
- Thompson, W. E. and Streib, G. F. (1958) 'Situational Determinants: Health and Economic Deprivation in Retirement', *Journal of Social Issues* 14(2): 18–34.
- Thompson, W. E., Streib, G. F., and Kosa, J. (1960) 'The Effect of Retirement on Personal Adjustment: A Panel Analysis', *Journal of Gerontology* 15(2): 165–9.
- Thurstone, L. L. (1927) 'A Law of Comparative Judgment', *Psychology Review* 34: 273–86.
- Thurstone, L. L. (1928) 'Attitudes Can Be Measured', *American Journal of Sociology* 33(4): 529–54.
- Thurstone, L. L. (1935) *The Vectors of Mind*. Chicago, IL: University of Chicago Press.
- Tissue, T. (1972) 'Another Look at Self-Rated Health Among the Elderly', *Journal of Gerontology* 27(1): 91–4.
- Tuckman, J. and Lorge, I. (1953) 'Somatic and Psychological Complaints of Older People in Institutions and at Home', *Geriatrics* 8: 274.
- Ware, J. E. Jr., Davis-Avery, A., and Donald, C. A. (1978) *Conceptualization and Measurement of Health for Adults in the Health Insurance Study: Vol. 5. General Health Perceptions*. Santa Monica, CA: Rand Corporation.
- Ware, J. E. Jr. and Sherbourne, C. D. (1992) 'The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection', *Medical Care* 30(6): 473–83.
- Wiggins, J. W. and Schoeck, H. (1961) 'A Profile of the Aging: U.S.A.', *Geriatrics* 16: 336–42.
- Woolsey, T. D. (1952) *Estimates of Disabling Illness Prevalence in the United States*. Washington, DC: Federal Security Agency, Public Health Service.
- Young, D. (1941) 'Memorandum on Suggestions for Research in the Field of Social Adjustment', *American Journal of Sociology* 46(6): 873–86.
- Young, J. L. (2017) 'Numbering the Mind: Questionnaires and the Attitudinal Public', *History of the Human Sciences* 30(4): 32–53.

Author biography

Tiago Moreira is Professor of Sociology at Durham University. His research focuses on the role of standardisation, quantification, and measurement in biomedicine, health care, and ageing, often using historical and ethnographic extended case studies. Recent publications include 'Translating Cell Biology of Ageing? On the Importance of Choreographing Knowledge' (2021) *New Genetics and Society* 40(3): 267–83.