The emergence of university rankings: a historical-sociological account

Stefan Wilbers[1] · Jelena Brankovic[1]

Abstract

Nowadays, university rankings are a familiar phenomenon in higher education all over the world. But how did rankings achieve this status? To address this question, we bring in a historical-sociological perspective and conceptualize rankings as a phenomenon in history. We focus on the United States and identify the emergence of a specific understanding of *organizational performance* in the postwar decades. We argue that the advent of this understanding constituted a discursive shift, which was made possible—most notably but not solely—by the rise of functionalism to the status of a dominant intellectual paradigm. The shift crystallized in the rankings of graduate departments, which were commissioned by the National Science Foundation and produced by the American Council on Education (ACE) in 1966 and 1970. Throughout the 1970s, social scientists became increasingly more interested in the methods and merits of ranking higher education institutions, in which they would explicitly refer to the ACE rankings. This was accompanied by a growing recognition, already in the 1970s, that rankings had a place and purpose in the higher education *system*—a trend that has continued into the present day.

Keywords Rankings · Universities · Performance · Institutionalization · Historical sociology · United States

Introduction

In accounting for the ubiquity of university rankings, scholars are overwhelmingly preoccupied with "successful" contemporary examples, such as the *U.S. News & World Report*, Shanghai Rankings or QS. The influence of these rankings is often attributed to broader trends, which are seen as impacting higher education over the recent decades, such as "marketization", "managerialism" and "neoliberalism". This has arguably diverted attention

---

away from the phenomenon of rankings itself and not least from the social and historical circumstances which have made its institutionalization possible.

In this article, we bring in a historical-sociological perspective and offer a corrective to these accounts. To this end, we conceptualize university rankings as a social operation whose legitimacy is rooted in a specific understanding of *organizational performance*—uniquely articulated by rankings. The crux of this understanding is the possibility to improve, which, as rankings become institutionalized, transitions into a widely shared belief that improvement is only possible in relation to the performance of other organizations (Brankovic et al., 2018). To trace the emergence of this understanding of performance in higher education, we analyse the historical evolution of university rankings in the twentieth century. Drawing inspiration from the recent work by Abend (2020), we ask the following what-makes-it-possible type of question: which broader historical conditions made the emergence of the rankings-specific idea of performance possible in the higher education context?

Empirically, we look at the changing discursive and institutional conditions in US higher education over the course of the twentieth century and in particular at the developments in the postwar decades. We have two reasons to focus on the United States and on this period in history. First, it is in this context that university rankings became widely regarded and consequential before they became so internationally, starting with the today well-known *U.S. News & World Report* ranking (USN), first published in 1983 (Myers & Robe, 2009; Sauder, 2008). Second, as far as the available evidence goes, for decades before the USN took the stage, scholars, administrators, national associations and even federal authorities would make frequent use of the ranking format to compare departments, colleges and universities, for various purposes and with varying effect (Hammarfelt et al., 2017; Webster, 1986).

We develop the argument in the following steps. We first lay down our understanding of what modern rankings are, conceptually speaking. In the second step, we show that throughout the first half of the twentieth century, the discourse on university comparison in the US largely revolved around standards and relatively stable classification systems. In the central part, we evidence a discursive shift, whereby universities are increasingly understood as *performing* entities in an increasingly more dynamic environment. This means that their respective performances were now (a) tied to a national higher education *system*; (b) established in *relation* with the performances of other universities in the system; and (c) seen as *continuously changing*. This shift took place roughly during the 1960s and 1970s, and it found a fitting expression in the already existing practice of *ranking*, which contributed to the further normalization of the idea that performances of higher education institutions could be plausibly quantified, compared and rank-ordered on a continuous basis. We conclude with a discussion on the implications and an outlook.

What does it mean to perform in a ranking?

We conceptualize rankings as quantified zero-sum comparisons of performances, visualized by means of a hierarchical table and repeatedly published by a third party (Werron & Ringel, 2017). Contemporary university rankings incorporate all of these elements whereby each is indispensable to their effect on the higher education field(s) (Brankovic et al., 2018). Considered individually, however, comparison of performances, quantification, hierarchical tables and publishing thereof are not unique to rankings nor did any of

these elements historically for the first time emerge in rankings (Hammarfelt et al., 2017; Ringel & Werron, 2020). Thus, accounting for the institutionalization of rankings in higher education necessarily leads to an empirical investigation into *when* and *how* each of the elements constituting them entered discourses[1] on academic quality and excellence and under what circumstances they merged to the form of ranking we are familiar with today.

Here it is of critical importance to distinguish between *performance in absolute terms*, which may be determined by reference to a standard, and *performance in relative terms*, which can only be established by comparing performances of two or more performers. Performing in a ranking is by definition of the relative kind, whereby the comparisons of performances are turned into a *zero-sum hierarchy*. This practically means that, in a ranking, the improvement in performances of one university comes inevitably at the expense of others. This understanding of performance is different from the one we find in other evaluation practices based on quantification and commensuration (Espeland & Stevens, 1998), such as ratings and benchmarks (cf. Klasik & Hutt, 2019), as well as those in various assessment schemes commonly found in higher education (cf. Hamann, 2016).[2]

In addition to being established by means of zero-sum comparisons, the performance in contemporary rankings also has a *temporal* dimension (Landahl, 2020; Ringel & Werron, 2021). The fact that rankings are published repeatedly and often regularly allows for a comparison of actors' performances over time. This likely first took off in the nineteenth-century sports (Eichberg, 1974; Parry, 2006). In a recent study, Minnetian and Werron (2021) show how the development of statistical practices, systematic data collection and the continuous production of tables over a season played a decisive role in the emergence of a new, *serial* understanding of achievement and performance in baseball. The idea that the performances of universities are compared *continuously*, from one ranking cycle to the next, is integral to university rankings as we know them today (Brankovic et al., 2018).

This understanding of performance both resonates with and reinforces the widely shared imaginary of higher education as a stratified order in which universities are expected to continuously perform as competitors for status (e.g. Brankovic, 2018; Cantwell & Taylor, 2013; Marginson, 2008). And yet, as we argue in this article, this understanding of performance was not born in 2003 with the first global rankings nor was it introduced by the USN rankings in 1983. Rather, its roots take us back to the early decades of the twentieth century.

Rankings in the United States: slow beginnings

A closer look at the literature on the history of higher education in the US does not suggest that rankings occupied an important place in the broader discourse on higher education during the first half of the twentieth century (Callahan, 1962; Geiger, 2005, 2014; Gelber, 2020). What we do find in the literature instead is a great deal of talk about standardization, classification and accreditation and particularly on what it meant to be a "real university" (Brubacher & Rudy, 1958; Veysey, 1965). Accordingly, the status orders of higher

---

[1] We use the terms "discourse" and "discursive" to broadly denote ways of thinking that we identify in the historical evidence we refer to.

[2] We also stress the *public* dimension of rankings. Many assessments which involve rank-ordering competitors do not make the zero-sum tables publicly available. This is the main reason why we do not consider these a ranking, even when they do produce stratifications.

education institutions largely revolved around conforming to standards, the major classification and accreditation schemes or membership in associations, such as the Association of American Universities (AAU) (Geiger, 2014; Hollis, 1938; Veysey, 1965).

During the Progressive Era, roughly between 1890 and 1920, a movement to "standardize" swept across societal domains in the US, including that of higher education (Geiger, 2005; Rudolph, 1962). As a result, organizational heterogeneity decreased, educational institutions cut their ties with religious denominations, the "old-time college" disappeared and new ideas of how to organize education and research emerged. Ideal models and standards for colleges, professional schools and an elaborated form of the "university", largely inspired by their European counterparts, were implemented by influential actors such as the AAU, the Carnegie Foundation for the Advancement of Teaching (CFAT) and several newly established professional organizations (Gelber, 2020; Hollis, 1938). If universities at that time were expected to perform as organizations, they would be expected to perform to a standard, not by entering competition with each other (Wilbers et al., 2021).

Although not rankings in the strict sense, the classification schemes emerging in this period are crucial to observe as they reveal the usual practices back in the day with respect to publicly evaluating and comparing higher education institutions in the US. It is therefore important to draw attention to the characteristics of some of the major studies conducted at the time that produced rank-ordered classes of universities. One distinctive characteristic was that they principally relied on the personal insight of an individual of some authority judging the quality of a university. For example, Charles Babcock—the author of one of the first classifications commissioned by the AAU—personally visited "nearly all of the large institutions having graduate schools", where he, among others, spoke with deans and presidents, examined student records and studied how they dealt with admissions (Babcock, 1911, p. 3). Similarly, in preparing his well-known study on medical schools, Abraham Flexner visited each of the 155 schools in the US and Canada included in his study (Flexner, 1910).

Second, the categories of comparison in these studies often served more for orientation than for exact comparison. For example, Flexner's evaluations varied across schools, and he took considerable freedom in the descriptions. The better the school, in his view, the shorter the comments, save for a few exceptionally good ones, which were to serve as models for the rest (Flexner, 1910). Journalist Edwin Slosson, the author of the widely read *Great American Universities* (Thelin, 1983; Veysey, 1965), offered mostly anecdotal evidence on what the success of a university was to be attributed to. He evaluated diverse things, such as the universities' architecture, how it made the visitor feel, the number of staff and students, the president's personality, the university town, the attitudes and culture of students and much more, sometimes all in one and the same paragraph (Slosson, 1910).

Third, even though they did include tables and numbers, these studies were largely based on narrative accounts. The reason for this, however, was not necessarily the lack of quantitative data; rather, it likely had more to do with a belief that numerical evidence was inferior to what could be offered by the thick descriptions provided by the expert personally visiting each university. Flexner's study was, in fact, commissioned in order to make up for what was considered to be limited quantitative information offered by tabulated comparisons (Wilbers et al., 2021). Slosson, in a similar vein, stressed that the quantitative comparisons of universities in his book were to be seen only as additional information, not as primary: "In presenting these diagrams and statistics I do not wish to be understood as giving them an exaggerated importance. The really important things are incommensurable and uncountable" (Slosson, 1910, p. 474).

The first comparisons of higher education institutions that were both quantified and zero-sum also appeared in the first decades of the twentieth century. They were produced by James McKeen Cattell, a well-known social psychologist and the long-time editor of *Science* (Hammarfelt et al., 2017; Ringel & Werron, 2020). Although Cattell eventually lost interest in reproducing the rankings, his work did inspire some of his contemporaries. Two notable examples were the biologist B. W. Kunkel and the education professor W. C. Eells, who would occasionally publish rankings of colleges and universities, sometimes even drawing on Cattell's methodology (e.g. Eells, 1926; Kunkel, 1924). These rankings would appear in *School and Society*, which was a periodical established in 1915 by the Society for the Advancement of Education and edited by Cattell for the next two decades (McInerny, 2008).

These quantified and tabulated comparisons of colleges and universities would appear more frequently during the interwar period and would be published in directory-style periodical publications, produced by, among others, the American Council on Education (ACE). An important driver of this trend was a growing interest and capacity for the systematic and continuous measurement of both student and faculty performances, in part driven by the business-inspired "cult of efficiency" (Callahan, 1962; Kliebard, 2004). Although these data would be used by institutions to compare achievements of individuals (Godin, 2009; Rudolph, 1962), public comparisons of thus aggregated data for organizations would be usually resisted by universities, in particular if they were to be produced by non-academic bodies, such as private foundations (Gelber, 2020).

The interwar period was also characterized by the increasingly more widespread use of standardized data collection methods, such as surveys, which were promoted both by the philanthropic foundations and the associations of universities and colleges (Eells, 1937). Although Cattell did use surveys to collect some of the data he used for his rankings (Hammarfelt et al., 2017), the first rankings based entirely on surveys were produced in 1925 by Raymond Hughes. Hughes, who was a chemistry professor, devised a reputation survey of graduate departments across the country which were "doing graduate work of some distinction", the results of which he turned into rankings (R. M. Hughes, 1925, p. 3). However, when he repeated the survey in 1934, this time as the chairman of the ACE, for some reason he refrained from ranking. Instead, he used the survey data to differentiate between the departments which were "distinguished" from those which were "adequately staffed and equipped" (R. M. Hughes, 1934). This hesitation to clearly rank-order the departments speaks to the still uncertain status of rankings at the time.

Even though Hughes based his analysis on standardized data collection and quantitative methods, large portions of his report, in particular in his 1934 study, were reserved for the names of the academics that completed the survey. Thus, although Hughes did not make personal visits to the institutions he studied, his choice to name the "jury", as he referred to those individuals, was indicative of the enduring importance of the individual credibility in judgements of institutional quality at the time. Nevertheless, the use of the survey method instead of narrative comparisons, and not least the peripheral nature of the author's personal judgement, placed Hughes' reports in contrast with the studies conducted by Babcock, Slosson or Flexner.

Hughes' studies were still criticized by his contemporaries, which, however, did not seem to lead to an updated version of his rankings; until in 1959 the task was picked by Hayward Keniston, a philology professor at the University of Pennsylvania (Lawrence & Green, 1980). Using Hughes' 1925 survey as a methodological inspiration, Keniston produced a ranking of 25 graduate departments belonging to the universities in the AAU. And

like Hughes, Keniston too was cautious about giving too much weight to the quantitative differences:

> Although the institutions have been ranked in accordance with the actual scores, the figures cannot be interpreted as having real validity. Differences of a few points are plainly meaningless. But there is sound reason to believe that those that are rated in the top five are really the outstanding departments. Neither the order nor the composition of the second five is as well established. And the third five is so uncertain that they might well have been omitted. They are included because they reveal the emergence of departments at institutions which do not have high over-all strength. (Keniston, 1959, p. 117)

It is interesting to note in this passage that Keniston's approach to the status differentiation seemed very much classificatory, even though it was presented as a zero-sum hierarchy. The fact that he made this explicit in the report speaks to the tension between these two approaches.

Keniston's decision to focus on a handful of elite institutions would soon be seen as unsuitable for a rapidly expanding higher education (Cartter, 1966). The developments which made such criticism increasingly more plausible were, in fact, already gaining momentum at the time Keniston published his rankings.

A discursive shift in the understanding of university performance

In this section, we focus on the discursive and institutional "enablers" (Abend, 2020), which made it possible for rankings to become a legitimate and an increasingly more used method of publicly comparing higher education institutions in the US. We do not aim to offer a definitive list of enabling conditions. Rather, we see each enabling condition as resting on further enabling conditions, and we only go as far in the chain of enablers as we find necessary for the argument. We also stress that we do not treat these enablers as causal explanations but as discursive and institutional conditions that played a critical role in the history of higher education rankings as we conceive of them in our conceptual framework. Finally, we intentionally do not write an exhaustive history of all relevant developments for the US higher education in the period covered, but we instead focus on those aspects which we find to be directly relevant for the history of rankings.

We focus on three discursive enablers, which were more or less specific to the US post-war context. The first enabler refers to the performance of a university being discursively *tied to the performance of the national higher education "system"*. The second enabler builds on the first, and it refers to the performance of one university being constructed as *relative* to the performances of other universities in the system. The third enabler, finally, refers to the performance of a university constructed as *continuously changing* and therefore established through repeated quantitative assessments and comparisons published by a third party.

In the following section, we examine the origins of these three discursive enablers and their institutionalization in the US higher education. Specifically, we highlight how each of them contributed to the new understanding of what it meant to perform as a higher education institution. In the subsequent section, we show how this understanding crystallized in rankings in the 1960s. Finally, we show how the practice of ranking was normalized

during the 1970s, further propelling the new understanding of performance in academic discourse.

The rise of the "system"

The idea of higher education as a "system" can be largely attributed to the rise of functionalism to the status of a dominant intellectual paradigm in the postwar period and its role in shaping both academic and policy discourses (Heyck, 2015). In a nutshell, functionalism called for every aspect of society, culture and behaviour to be defined and examined in systemic terms. This meant, roughly put, seeing the social world as a complex, hierarchical structure, made of interrelated units and subunits, functions and processes, in which action was purposeful and behaviour adaptive. By offering a highly rationalistic vision of society, functionalism presented itself as a fitting way of thinking also for the increasingly more technocratic style of federal policy making (Heyck, 2015).

Together with its intellectual derivative, modernization theory, functionalism had a profound impact on the national policy discourse in the postwar decades (Cohen-Cole, 2014; Gilman, 2003; Heyck, 2015; Jardini, 2000). Federal planning, which had been present in some areas already in the period of the New Deal, became a common macro-policy approach across sectors during the 1950s and 1960s. This development was accompanied by an increased interest in examining complex issues and in devising ever more efficient ways of addressing them. As a result, between the second half of the 1950s and in the 1960s, the amount of data collected and studies produced in virtually every policy area was unprecedented. This was particularly enabled by the advancements in computer technology and increasingly more sophisticated methods and tools for data collection and analysis (Astin, 2003; Hutt, 2017).

Higher education institutions—as a home of both education and science—were of strategic interest to the federal authorities during this period. This interest is usually interpreted against the backdrop of the Cold War, with standard reference to the events such as the launching of Sputnik and Gagarin's space journey (Geiger, 1997; Wolfe, 2013). These events were important in at least two ways. First, they instilled in policy makers and university representatives a view of the American education and science as a collective national project. Second, they led to a significant increase in federal funding for both university education and university-based research. On the research side, the federal authorities were especially interested in expanding capacities by investing in state-of-the art scientific facilities on university campuses across the country. Accompanied by a dramatic increase in student enrollments in the 1960s, this "federal largess" led to "an ephemeral golden age in American higher education" (Geiger, 2005, p. 65).

It was during this period that the appreciation of higher education as a system came of age.[3] The functionalist "system thinking" also informed policy discourse, not least by shaping the outlook of some of the most prominent intellectual figures in the policy scene in the 1960s and 1970s, such as Clark Kerr and Martin Trow, as well as many of their peers among university administrators and scholars. Kerr's work stood out in particular,

---

[3] We do not claim that the phrases such as "higher education system", "(American) system of higher education", "college system" or "university system" were first used in the mid-twentieth century. In fact, invoking a "system" was not uncommon even in the nineteenth century. However, its use was often casual, in addition to lacking a specific understanding of performance relations of interdependent parts, which was now associated with the term "system".

especially during the 1960s and 1970s, first as the leader of the California Master Plan for Higher Education and then as the chair of the Carnegie Commission on Higher Education and the Carnegie Council on Policy Studies in Higher Education (Marginson, 2016; Wellmon, 2021). Referring to the historical "marriage" of functionalism and higher education in this period, Wittrock (1993) would write some decades later, "functionalism gave American university representatives a self-understanding which seemed to make perfect sense of the realities of their institutional situation" (p. 337). In light of this, Kerr's California Master Plan of 1960, which inspired blueprint planning in other parts of the US (Marginson, 2016), was emblematic of the "system thinking" about the plurality of roles higher education institutions played in society.

It was perhaps not a coincidence that Talcott Parsons—a leading figure in systems theory and functionalist thought—would often use higher education and science as an empirical subject in his publications (Parsons & Platt, 1968, 1973). For Parsons and Platt, no other academic system in the world "remotely approximates the peaks reached by the most eminent American universities" (Parsons & Platt, 1968, p. 497). This superiority in performance, they argued, could be attributed to a value system of "instrumental activism", a kind of specifically American Protestant ethic that drives people to constantly strive to improve, bringing about a productive system of closely interrelated universities as its potent parts (Parsons & Platt, 1968). Similar reflections and further considerations on the connection between the university "system", on the one hand, and "performance" and "achievement", on the other, can be found in numerous other authors, which often directly drew on the writings of Parsons and colleagues.

The idea of universities as always oriented towards improvement pushed forward an understanding of excellence as attainable through performance. Although the idea that universities *could perform into* or *achieve* excellence by performing well in specific domains was not in itself new during the 1950s, it was during this period that it became explicitly recognized and promoted in national discourse. John W. Gardner, the president of the CFAT, discussed the notion of excellence in his widely acclaimed book *Excellence: Can We Be Equal and Excellent Too?* (1961). Referring to higher education institutions, he wrote:

> We do not want all institutions to be alike. We want institutions to develop their individualities and to keep those individualities. None must be ashamed of its distinctive features so long as it is doing something that *contributes importantly to the total pattern*, and so long as it is *striving for excellence in performance*. (Gardner, 1961, p. 83, emphasis added).

Gardner's emphasis on "excellence in performance" resonated with the argument Robert K. Merton's made at the time, that turning "potential excellence" into "actuality" was something to be pursued (Merton, 1973, p. 425 [1960]).[4] And while achievement had been long recognized as worth of encouraging in individuals, the "system thinking" allowed for its extension to include also other units of the system, such as organizations, as well as the system as a whole.

---

[4] This thinking also resonated with the idea of "meritocracy" which emerged around that time, coined by Michael Young in his widely read book, *The Rise of the Meritocracy* (1958). He used the term to refer to a fictional (British) society, set in the future, whose ruling class was determined not by inheritance, but by a formula "IQ + effort = merit".

This notion of *achievement* was particularly popularized by the proponents of then emerging modernization theory (Gilman, 2003; Knöbl, 2003). In their effort to explain macro-phenomena such as "modernity" and "progress", prominent intellectual figures, such as Walt Whitman Rostow, Marion Levy and David C. McClelland, put forward the distinction between "traditional" or "pre-modern" and "industrial" "modern" or in McClelland's case "achieving societies". In this view, and put in Parsonian terms, traditional societies would judge actors for their ascribed qualities, while modern societies would base their judgements on performances or achievements. Achievement was thus something that could be defined as "success in competition with some standard of excellence" (McClelland et al., 1953, p. 110) as well as something that could be planned and organized through instrumental reason.

What did it mean for higher education institutions to *achieve* "excellence in performance"? For some, this meant embracing an entrepreneurial vision and a strategic approach to the acquisition of federal funding and industrial patronage (Lowen, 1997). For others, it meant offering doctoral education (Berelson, 1960). Bernard Berelson, a behavioural scientist and a programme officer at the Ford Foundation, found the ambition of graduate departments and universities across the country to "get ahead" and "climb" as natural and something to be encouraged:

> Just as the way for the academic man to get ahead was to earn the doctorate, the way for an institution to get ahead was to offer it. "There is no man who does not need to climb." Neither, apparently, are there many institutions—and in our educational system, climbing means getting into the big league of graduate, and especially doctoral, study. (Berelson, 1960, p. 135)[5]

This imperative to "rise" stood in contrast with thinking about status differences between universities as relatively stable over longer periods and from judging their performances on the basis of time-honoured standards, which had shaped the higher education discourse in earlier periods. And although many of the universities in Berelson's "big league" had been among Slosson's "great American universities", Berelson found it important to stress that any university, if it performed well enough, could "get ahead" and "get into" the big league of research universities.

There were two important implications of explicitly encouraging higher education institutions to strategically strive for excellence in performance. First, it suggested that performances could be observed over time, allowing for the present-moment actions to be plausibly tied to both past and future achievements. Second, it invited a new kind of understanding of the higher education institutions' own role and place in the "national system", which simultaneously opened up the possibility for the actors with an interest in the overall system performance, such as the National Science Foundation (NSF), to gain in prominence.

The ACE rankings of graduate departments

The increase in science funding in the post-Sputnik period brought the federal science bodies, such as the NSF, to the forefront of the national higher education scene. Faced with the

---

[5] The idea of getting ahead by emulating higher-status institutions was also famously captured with Riesman's earlier metaphor of the "snakelike procession" (1956).

expansion of graduate departments across the country and a pressure to allocate a growing amount of research money to a maximum effect, the NSF was on the look for a new way of deciding on graduate programme funding. "NSF's basic selection mechanism", reported Allan M. Cartter a decade later, "had been the peer-group judgment of proposals, but they wanted to make it more efficient, sophisticated, and presumably more extensive" (Dolan, 1976, p. 26).

In search for a solution, in 1963, the NSF officials met with the ACE president and vice president, Logan Wilson and Allan M. Cartter. On this occasion, Richard Bolt, the NSF's assistant director, informed Wilson and Cartter that the NSF had been considering to commission an evaluation and ranking of science departments to a private research group (Dolan, 1976). The reaction of the ACE's officials to this idea is worth reporting in full as it encapsulates the position universities had held at the time about the prospect of being evaluated by "outsiders". In Cartter's own words:

> Logan Wilson and I hit the ceiling and alerted him (Bolt) to the fact that there had been a long-standing position of the Association of Graduate Deans that nobody ought to play around with evaluating the graduate programs who was not themselves responsible to the institutions themselves; either the AAU or some other group. The deans had always said that anyone who is going to accredit us or rate us has to be responsible to the institutions. We don't want any outside body doing this. (Dolan, 1976, p. 26)

A compromise was, however, found: the ACE offered to produce the ranking by surveying the graduate departments itself. Cartter, at the time also the director of a newly established ACE Commission on Plans and Objectives for Higher Education, would be the one responsible for the task. The study was financially supported by the NSF, the National Institute of Health and the Office of Education. The contribution of the Office of Education was about one-quarter of the total sum, and its inclusion was important in order to "show that it was not strictly a 'scientific' interest group". At the same time, it was not supposed to be too big a contribution so as not "to alarm the institutions with what might be called a 'federal' ranking" (Dolan, 1976, p. 27). Even though by the early 1960s, leading university administrators viewed federal patronage as a necessity, and not any more as a threat (Lowen, 1997), the ACE's approach suggests that federal involvement was a sensitive matter.

The rankings of graduate departments, which were the central part of what was later called *The Cartter Report* (1966), were based on a large-scale quantitative reputation survey. They were a compromise between the NSF's desire for an "objective" and "scientifically sound" performance assessment and the ACE's desire to insulate the performance comparisons from the "outsiders". Notably, Cartter explicitly aimed for the rankings to build on the earlier works in the genre, namely, those produced by Hughes and Keniston. But he also found it important to expand, test and corroborate his survey data with a wealth of additional data and rigorous statistical methods, which he extensively discussed in the report. In this sense, unlike Keniston and Hughes, who were more interested in qualitative properties *behind* the numbers, Cartter seemed to have believed in the power of numbers to *show* things which were otherwise not visible.

The foreword to the *Cartter Report*, penned by Logan Wilson, opened with the following statement:

> EXCELLENCE, by definition, is a state only the few rather than the many can attain. Striving for academic excellence, however, is a worthy ideal for colleges and univer-

sities, and it can be reasonably argued that every educational institution should meet minimum qualitative standards, and particularly if it offers graduate work. A present problem is the need for a better general understanding of what quality signifies. (Wilson, 1966, p. vii, capital letters in original)

The central concern of the report, according to Wilson, was "what quality signifies". Yet irrespective of how one defines quality, he continued, "academic excellence" was considered a "worthy ideal" that all colleges and universities should *strive for*, which resonated with the earlier words by Berelson and Gardner. Minimum standards of quality were to be met by every institution, but "excellence" meant going beyond standards.

It is noteworthy that Cartter was acutely concerned with the uneven quality of universities across the country, in particular those located in the southern states: "In an age of national rather than regional competition—for faculty, students, foundation support and government contracts—Southern higher education *must become quality conscious or be left behind*" (Cartter, 1965, p. 69, emphasis added). The problem of the "lag in quality" of universities in the south in comparison with other parts of the country was considered in view of its potential consequences for the performance of "the economy" of those regions (Cartter, 1965).

As an economist specializing in, among other areas, "manpower planning", Cartter was interested in predicting trends in education and the labour market, which was a likely reason why a future outlook, accompanied by a sense of urgency, permeated much of his writing. The closing paragraph of the ACE report is illustrative in this regard:

Evaluative information, such as that presented in this study, will be little more than a curiosity unless the stronger graduate schools and the associations representing university presidents and deans take the initiative both in setting standards and in helping the smaller and weaker institutions to live up to those standards. It is hoped that this and successive surveys to be undertaken at approximately five-year intervals will be of value to these groups in their attempts to strengthen graduate education and thereby to invigorate all of American higher education. (Cartter, 1966, p. 121)

We note here that Cartter did not see universities as isolated entities, each catering to its own destiny; rather, they were there to "help" each other on their *common path* of invigorating "all of American higher education". This ambition to *see* the graduate education across the country as comprehensively as possible,[6] as well as in finding ways to identify which universities "hold *the most promise of improving* their relative positions in the years immediately ahead" (1966, p. 117 emphasis added), echoes the earlier described "system" approach in the understanding of university performances. This placed Cartter in contrast with Keniston and Hughes. Moreover, Cartter's view of performance as something to be *evidenced* by undertaking "successive surveys" at regular intervals made this contrast even sharper.

The *Cartter Report* sold in 26,000 copies, which was a significant circulation for a publication of this kind and was largely met with critical acclaim (Webster, 1986). The success was, however, not repeated by the ACE follow-up rankings, published in the *Roose-Andersen Report* (1970), even though its release had been eagerly awaited by the

---

[6] Cartter's rankings were explicitly meant to be "the most comprehensive ever undertaken" (Wilson 1966, p. vii), which they likely were, as they included 106 universities, 30 disciplines and around 4000 respondents to the reputation survey on which they were based.

administrators across the country (Dolan, 1976). Although Cartter had envisioned successive ACE rankings to be published at regular intervals, this eventually did not happen.[7] The NSF was probably less interested in serializing the rankings, while the ACE's focus shifted away from identifying potential excellence in graduate departments. And in fact, already in the *Roose-Andersen Report*, the emphasis was not as much on identifying (potential) excellence, as much as it was "to protect the potential consumer of graduate education from inadequate programs" (Roose & Andersen, 1970, p. 2), thus stressing the importance of minimum "adequacy" standard. This focus on potential "consumers", although not entirely a new idea in itself, reflected an already changing discourse and foreshadowed the ascendance of the commercial rankings in the 1980s. However, as we will show in the next section, the fact that no rankings on a par with the ACE ones were produced during the 1970s does not mean that the ACE rankings had no relevance for the further career of the practice, quite on the contrary.

The normalization of ranking as a social scientific method and practice

Particular to the 1960s and consistent with the expansion of higher education and science that marked the decade, there was a growing interest among scholars in measuring things like the "output", "productivity" or "prestige" of departments and institutions. This interest led scholars not only to directly refer to the ACE rankings, but also to produce own rankings, often drawing inspiration from the ACE reports.[8] These rankings would be published in academic journals, mostly in social sciences, although scholars in natural sciences would occasionally do the same in their own disciplinary outlets. By and large, to these scholars, the ACE reports were works of scholarship, and ranking was seen as a legitimate social scientific method and practice.

These studies drew on the ACE reports in various ways. For some, the ACE reports were a source of data, for others, a methodological inspiration. Some would offer suggestions for improving the method of ranking (e.g. Drew & Karpf, 1975), others would produce rankings of disciplines not included in the ACE reports (e.g. Carpenter & Carpenter, 1970; Edwards & Barker, 1977) or simply offer an alternative ranking of departments in their disciplines (e.g. Knudsen & Vaughan, 1969). Rankings would occasionally also be recognized as a solution for various problems. For instance, Glenn and Villemez (1970) saw sociology departments suffering from a general lack of "rivalry", and their ranking would, they hoped, "help promote rivalry among departments and aid prospective students in their choice of departments" and further "help promote the development of the discipline" (p. 244).

These developments made way for an emerging tradition in social sciences which revolved around discussing data, methodologies and general approaches to comparing and evaluating departments and universities across the country, also beyond graduate programmes. In general, among the social science disciplines, sociological journals stood out, although more specialized journals, monographs and edited volumes were not far behind.

---

[7] In 1982, the National Research Council published *An Assessment of Research-Doctorate Programs in the United States*, which was based on the ACE reports and also included rankings by disciplines.

[8] We retrieved from Google Scholar in total 1215 unique works published throughout the 1970s and 1980s which cited the *Cartter Report* and/or *Roose-Andersen Report*. From the total of 381 works citing both ACE reports, 309 were published in academic journals, out of which 77 in the field of higher education, 53 in sociology, 40 in economics and 36 in psychology.

Higher education journals also showed an interest in publishing rankings of departments and universities (e.g. Elton & Rodgers, 1973; Magoun, 1966; Margulies & Blau, 1973). These studies were frequently accompanied by the discussions on quality and excellence of academic organizations and debates on how they could be measured. Criticism explicitly directed at rankings also became more common throughout the 1970s (Dolan, 1976), which attests to the increasing interest in the ranking approach to comparing academic organizations.

The practice of producing and debating rankings among academics themselves would practically become self-sustained during the 1970s, which contributed to the "domestication" of the rankings-specific understanding of organizational performance in scholarly discourse. This is, of course, not to say that rankings produced by the social scientists in the 1970s were as consequential as they became later on. Rather, we argue that they contributed to the normalization of the idea that the performances of higher education institutions could plausibly be quantified, compared and rank-ordered on a continuous basis. The normalization of this understanding of performance meant that rankings could, in potential at least, be used as a legitimate source of information on the relative quality of higher education institutions both by the "insider" and "outsider" audiences, including prospective students.

When the *U.S. News & World Report* published its first rankings in 1983, being ranked—be it for reputation, productivity or selectivity—was far from a new experience, not only for graduate departments, but also for entire institutions, including undergraduate programmes. Crucially, the USN capitalized on a now already established understanding of performance—as zero-sum. In this sense, despite being an "outsider" to the field,[9] the USN rankings spoke to the concerns administrators and academics themselves already had, and it did this in a "language" in which they had already been fluent—the one of performance and rankings. But perhaps more importantly, it spoke to the realities of the expanding student market and the concerns both prospective students and higher education administrators had about the prestige of individual institutions in the national context. At the same time, the USN did bring a number of important novelties into the practice of ranking itself. First, it introduced its already extensive expertise from consumer markets research into higher education (Espeland & Sauder, 2016). Second, by repeating the rankings, the USN eventually institutionalized the seriality—something that Cartter had hoped to do but did not live long enough to see. In doing so, USN effectively exported the rankings-specific idea of performance—which was now increasingly more seen as a proxy for quality and excellence in academic organizations—to a much wider audience.

Conclusion and outlook

We started this article with the premise that conceptualizing academic rankings as a social phenomenon that evolved in history would allow us to expose the cultural underpinnings of the rankings' ubiquity. By conceptualizing rankings as, among others, comparisons of performances, we set out to unravel the historical circumstances under which rankings became a legitimate method of comparing higher education institutions based on continuously

---

[9] USN was neither the first non-academic higher education ranker. *The Gourman Report* had been around since the 1960s, first rating, and then later in the 1970s, also ranking undergraduate programmes (Webster 1984).

published observations. This, however, does not mean that standards and classifications disappeared. The widely referred to Carnegie Classification of Institutions of Higher Education, for example, was conceived around the same time as the ACE rankings and has also continued into the present. It should neither stay unmentioned that, when USN developed its first rankings, it in fact relied on the Carnegie Classification regarding which institutions to include in its reputation survey (Solorzano & Quick, 1983). Despite the ascendance of rankings, classifications have remained an important part of the institutional landscape, in the US as well as elsewhere.

Functionalism and modernization theory may have long been proclaimed dead, but their discursive legacy continues to live through the understanding of higher education as a "system" and of universities as continuously performing entities. The federal planning in its former largess has also been abandoned almost completely, while neither the NSF nor the ACE retained a particular interest in continuing to fund or publish academic rankings. And in fact, not long after, a new set of enablers would emerge in the form of the popular press working their way to address the needs of what was perceived as a promising new market. Finally, the sociologists of science may have lost interest in publishing rankings in their own journals, but the scholarly interest in debating the methodological merits of various rankings has been kept alive in, among others, the academic fields of higher education and science studies. In this sense, rankings became increasingly more recognized as meaningful and therefore necessary to do "properly". It was this recognition that kicked off a historical quest for ever-better ways of ranking higher education institutions, which we can still witness today. When rankings hit the world stage in the early 2000s, all of these elements came together. The scale and the actors may have been different, but the logic of the social operation constituting rankings was—and continues to be—largely the same.

We see the following implications of our findings for the general study of rankings in higher education. First, accounting for the institutionalization of rankings requires both a theoretically grounded understanding of what rankings are as a phenomenon and an understanding that history is more than an inventory of events and individuals. Our approach in this article has been historical-sociological, which led us to treat the discursive realm as the prime site for an empirical investigation into what we have herewith referred to as institutional and discursive enablers. This implies that, while the rankings produced by, for example, Cartter may have been in their material aspects very similar to those Hughes produced several decades earlier, the meanings they were given in their respective historical contexts, and not least the historical contexts themselves, were essentially very different. At the same time, these two rankings were not unrelated: it is possible that Cartter's approach to the task of ranking would have been different had there not been for the earlier rankings by Hughes or Keniston. Therefore, the historical-sociological inquiry into why a particular ranking today is the way it is means identifying trajectories that are specific to its emergence and evolution.

Second, approaching rankings as a social and historical phenomenon, which has evolved in parallel to and as a part of other phenomena, urges us to look beyond specific instances, such as the USN rankings or the *Cartter Report*. But it also urges us to look beyond the short-sighted diagnoses such as that rankings are to be causally and casually attributed to unspecific phenomena such as "neoliberalism". In this sense, we agree with Hammarfelt et al. (2017) in the observation that "the popularity of university rankings cannot solely be explained by increasing top-down governance in neoliberal academia, because the practice of ranking ties in with deeply engrained cultural repertoires around competition and performance" (p. 392). Our approach takes the observation a step further as it traces the historical development of the said cultural repertoires in higher

education. Future research could thus explore the historical entanglement of rankings and other developments specific or non-specific to particular national and international historical circumstances beyond the US.

Third, by definition, rankings are accompanied by a specific understanding of what it means to perform as a university. Therefore, as rankings become institutionalized in discourse and policy, so does the understanding of organizational performances as improvable exclusively though entering competition with other universities. Clearly, there are other ways to evaluate and compare performances of universities, and these often exist alongside rankings. However, given the dramatic proliferation of all kinds of rankings over the last several decades, and not least their increasing deployment for various purposes, it is not unlikely that their diffusion has led to the displacement of other methods of university performance evaluation in specific contexts. This is something that should be empirically investigated.

Finally, the approach we have put forward with this article sensitizes our analytical eye for the ever-changing broader discursive and institutional conditions—across national and international contexts—which make these changes possible. And while saying that "rankings are here to stay" would likely resonate with many, we would like to propose another way of looking at it, which follows from our study: "rankings are here to change and their status challenged". At the moment of writing these lines, we are approaching the end of the second decade of university rankings as a truly global phenomenon. Since then, the discourse on academic rankings has evolved, as have rankings themselves. Rankings have not just diffused across the globe, but they have also become more varied and more complex, and, crucially, they have become deeply embedded in the epistemic fabric of higher education and as such normalized. This, we contend, makes studying them a rather challenging task but—for the very same reason—a critical one to pursue.

Declarations

Conflict of interest The authors declare no competing interests.

References

Abend, G. (2020). Making things possible. *Sociological Methods & Research*. https://doi.org/10.1177/0049124120926204

Astin, A. W. (2003). From number crunching to spirituality. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (pp. 1–56). Dordrecht: Springer.

Babcock, K. C. (1911). *A classification of universities and colleges with reference to bachelor's degrees*. Washington, DC: Bureau of Education.

Berelson, B. (1960). *Graduate education in the United States*. New York, NY: McGraw-Hill.

Brankovic, J. (2018). The status games they play: Unpacking the dynamics of organisational status competition in higher education. *Higher Education, 75*(4), 695–709.

Brankovic, J., Ringel, L., & Werron, T. (2018). How rankings produce competition: The case of global university rankings. *Zeitschrift für Soziologie, 47*(4), 270–288. https://doi.org/10.1515/zfsoz-2018-0118

Brubacher, J. S., & Rudy, W. (1958). *Higher education in transition: An American history, 1636–1956*. New York, NY: Harper & Brothers.

Callahan, R. E. (1962). *Education and the cult of efficiency*. Chicago, IL: The University of Chicago Press.

Cantwell, B., & Taylor, B. J. (2013). Global status, intra-institutional stratification and organizational segmentation: A time-dynamic tobit analysis of ARWU position among U.S. universities. *Minerva, 51*(2), 195–223. https://doi.org/10.1007/s11024-013-9228-8

Carpenter, R. L., & Carpenter, P. A. (1970). The doctorate in librarianship and an assessment of graduate library education. *Journal of Education for Librarianship, 11*(1), 3–45. https://doi.org/10.2307/40322097

Cartter, A. M. (1965). Qualitative aspects of southern university education. *Southern Economic Journal, 32*(1), 39–69. https://doi.org/10.2307/1055924

Cartter, A. M. (1966). *An assessment of quality in graduate education*. Washington, DC: American Council on Education.

Cohen-Cole, J. (2014). *The open mind: Cold War politics and the sciences of human nature*. Chicago, IL: The University of Chicago Press.

Dolan, W. P. (1976). *The ranking game: The power of the academic elite*. Lincoln, NE: University of Nebraska Printing Service.

Drew, D. E., & Karpf, R. S. (1975). *Evaluating science departments: A new index*. Santa Monica, CA: RAND Corporation.

Edwards, R., & Barker, L. (1977). A rating of doctoral programs in speech communication, 1976. *Association for Communication Administration Bulletin, 20*, 59–69.

Eells, W. C. (1926). Educational research and statistics. *School and Society, 23*(591), 535–536.

Eells, W. C. (1937). *Surveys of American higher education*. New York, NY: CFAT.

Eichberg, H. (1974). Der Beginn des modernen Leistens. *Sportwissenschaft, 4*(1), 21–48.

Elton, C. F., & Rodgers, S. A. (1973). The departmental rating game: Measure of quantity or quality? *Higher Education, 2*(4), 439–446.

Espeland, W. N., & Sauder, M. (2016). *Engines of anxiety: Academic rankings, reputation, and accountability*. New York, NY: Russell Sage Foundation.

Espeland, W. N., & Stevens, M. L. (1998). Commensuration as a social process. *Annual Review of Sociology, 24*(1), 313–343. https://doi.org/10.1146/annurev.soc.24.1.313

Flexner, A. (1910). *Medical education in the United States and Canada: A report to the Carnegie Foundation for the advancement of teaching*. New York, NY: Merrymount Press.

Gardner, J. W. (1961). *Excellence: Can we be equal and excellent too?* New York, NY: Harper & Row Publishers.

Geiger, R. L. (1997). What happened after Sputnik? Shaping university research in the United States. *Minerva, 35*(4), 349–367.

Geiger, R. L. (2005). The ten generations of American higher education. In P. G. Altbach, R. O. Berdahl, & P. J. Gumport (Eds.), *American higher education in the twenty-first century: Social, political, and economic challenges* (2nd ed., pp. 38–70). Baltimore, MD: JHU Press.

Geiger, R. L. (2014). *The history of American higher education: Learning and culture from the founding to World War II*. Princeton, NJ: Princeton University Press.

Gelber, S. M. (2020). *Grading the college: A history of evaluating teaching and learning*. Baltimore, MD: JHU Press.

Gilman, N. (2003). *Mandarins of the future: Modernization theory in Cold War America*. Baltimore, MD: JHU Press.

Glenn, N. D., & Villemez, W. (1970). The productivity of sociologists at 45 American universities. *The American Sociologist, 5*(3), 244–252.

Godin, B. (2009). The value of science: Changing conceptions of scientific productivity, 1869 to circa 1970. *Social Science Information, 48*(4), 547–586. https://doi.org/10.1177/0539018409344475

Hamann, J. (2016). The visible hand of research performance assessment. *Higher Education, 72*(6), 761–779. https://doi.org/10.1007/s10734-015-9974-7

Hammarfelt, B., de Rijcke, S., & Wouters, P. (2017). From eminent men to excellent universities: University rankings as calculative devices. *Minerva, 55*(4), 391–411. https://doi.org/10.1007/s11024-017-9329-x

Heyck, H. (2015). *Age of system: Understanding the development of modern social science*. Baltimore, MD: JHU Press.

Hollis, E. V. (1938). *Philanthropic foundations and higher education*. New York, NY: Columbia University Press.

Hughes, R. M. (1925). *A study of the graduate schools of America*. Miami, FL: Miami University.

Hughes, R. M. (1934). Report of the committee on graduate instruction. *Educational Record, 15*(2), 192–234.

Hutt, E. L. (2017). "Seeing like a state" in the postwar era: The Coleman report, longitudinal datasets, and the measurement of human capital. *History of Education Quarterly, 57*(4), 615–625. https://doi.org/10.1017/heq.2017.36

Jardini, D. R. (2000). Out of the Blue Yonder: The transfer of systems thinking from the pentagon to the great society, 1961–1965. In A. C. Hughes & T. P. Hughes (Eds.), *Systems, experts, and computers: The systems approach in management and engineering, World War II and After* (pp. 33–461). Boston, MA: MIT Press.

Keniston, H. (1959). *Graduate study and research in the arts and sciences at the University of Pennsylvania*. Philadelphia, PA: University of Pennsylvania Press.

Klasik, D., & Hutt, E. L. (2019). Bobbing for bad apples: Accreditation, quantitative performance measures, and the identification of low-performing colleges. *The Journal of Higher Education, 90*(3), 427–461. https://doi.org/10.1080/00221546.2018.1512804

Kliebard, H. M. (2004). *The struggle for the American curriculum, 1893–1958* (3rd ed.). New York, NY: Routledge.

Knöbl, W. (2003). Theories that won't pass away: The never-ending story of modernization theory. In G. Delanty & E. Isin (Eds.), *Handbook of Historical Sociology* (pp. 96-107). London: SAGE.

Knudsen, D. D., & Vaughan, T. R. (1969). Quality in graduate education: A re-evaluation of the rankings of sociology departments in the Cartter Report. *The American Sociologist, 4*(1), 12–19.

Kunkel, B. W. (1924). The colleges and scientific leadership. *School and Society, XIX,* 411–422.

Landahl, J. (2020). The PISA calendar: Temporal governance and international large-scale assessments. *Educational Philosophy and Theory, 52*(6), 625–639. https://doi.org/10.1080/00131857.2020.1731686

Lawrence, J. K., & Green, K. C. (1980). *A question of quality: The higher education ratings game*. Washington, DC: AAHE-ERIC.

Lowen, R. S. (1997). *Creating the Cold War university: The transformation of Stanford*. Oakland, CA: University of California Press.

Magoun, H. W. (1966). The Cartter Report on quality in graduate education: Institutional and divisional standings compiled from the report. *The Journal of Higher Education, 37*(9), 481–492. https://doi.org/10.2307/1979136

Marginson, S. (2008). Global field and global imagining: Bourdieu and worldwide higher education. *British Journal of Sociology of Education, 29*(3), 303–315. https://doi.org/10.1080/01425690801966386

Marginson, S. (2016). *The dream is over*. Berkeley, CA: University of California Press.

Margulies, R. Z., & Blau, P. M. (1973). The pecking order of the elite: America's leading professional schools. *Change, 5*(9), 21–27.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. (1953). *The achievement motive*. New York, NY: Appleton-Century-Crofts.

McInerny, P. M. (2008). Nicholas Murray Butler, James McKeen Cattell, and the Educational Review: Footnote to a famous feud. In R. L. Geiger (Ed.), *Perspectives on the History of Higher Education* (pp. 129–146). London: Routledge.

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. Chicago, IL: The University of Chicago Press.

Minnetian, C., & Werron, T. (2021). Redefining achievement: The emergence of rankings in American baseball. In L. Ringel, W. Espeland, M. Sauder, & T. Werron (Eds.), *Worlds of rankings* (pp. 127–151). Bingley: Emerald Publishing Limited. https://doi.org/10.1108/S0733-558X20210000074031.

Myers, L., & Robe, J. (2009). *College rankings: History, criticism and reform*. Washington, DC: Center for College Affordability and Productivity.

Parry, J. (2006). The idea of the record. *Sport in History, 26*(2), 197–214. https://doi.org/10.1080/17460260600786864

Parsons, T., & Platt, G. M. (1968). Considerations on the American academic system. *Minerva, 6*(4), 497–523. https://doi.org/10.1007/BF01096545

Parsons, T., & Platt, G. M. (1973). *The American university*. Cambridge, MA: Harvard University Press.

Riesman, D. (1956). *Constraint and variety in American education*. Lincoln, NE: University of Nebraska Press.

Ringel, L., & Werron, T. (2020). Where do rankings come from? A historical-sociological perspective on the history of modern rankings. In A. Epple, W. Erhart, & J. Grave (Eds.), *Practices of comparing: Towards a new understanding of a fundamental human practice* (pp. 137–170). Bielefeld: Bielefeld University Press.

Ringel, L., & Werron, T. (2021). Serielle Vergleiche: Zum Unterschied, den Wiederholung macht. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie, 73*(1), 301–331. https://doi.org/10.1007/s11577-021-00750-2

Roose, K. D., & Andersen, C. J. (1970). *A rating of graduate programs*. Washington, DC: American Council on Education.

Rudolph, F. (1962). *The American college and university: A history*. Athens, GA: University of Georgia Press.

Sauder, M. (2008). Interlopers and field change: The entry of U.S. news into the field of legal education. *Administrative Science Quarterly, 53*(2), 209–234. https://doi.org/10.2189/asqu.53.2.209

Slosson, E. E. (1910). *Great American universities*. New York, NY: Macmillan.

Solorzano, L., & Quick, B. E. (1983). Rating the colleges: Exclusive national survey. *U.S. News & World Report, 41*(41–46), 48.

Thelin, J. R. (1983). Once and almost great American universities? *The Review of Higher Education, 6*(2), 161–168. https://doi.org/10.1353/rhe.1983.0028

Veysey, L. R. (1965). *The emergence of the American university*. Chicago, IL: The University of Chicago Press.

Webster, D. S. (1984). Who is Jack Gourman: And why is he saying all those things about my college? *Change: The Magazine of Higher Learning, 16*(8), 14–19.

Webster, D. S. (1986). *Academic quality rankings of American colleges and universities*. Springfield, IL: Charles C. Thomas.

Wellmon, C. (2021, April 22). The crushing contradictions of the American university. *The Chronicle of Higher Education*. https://www.chronicle.com/article/the-crushing-contradictions-of-the-american-university. Accessed 14 June 2021

Werron, T., & Ringel, L. (2017). Rankings in a comparative perspective. Conceptual remarks. In S. Lessenich (Ed.), *Geschlossene Gesellschaften. Verhandlungen des 38. Kongresses der Deutschen Gesellschaft für Soziologie in Bamberg 2016* (pp. 1–10). Essen: DGS.

Wilbers, S., Ringel, L., & Werron, T. (2021). Homöopathen, 'Quacksalber' und wissenschaftliche Mediziner. Zu den Anfängen der Hochschulrankings in der medizinischen Ausbildung der USA, 1850–1930. In F. Meier & T. Peetz (Eds.), *Organisation und Bewertung* (pp. 393–424). Wiesbaden: Springer VS.

Wilson, L. (1966). Foreword. In A. M. Cartter (Ed.), *An assessment of quality in graduate education* (pp. vii–viii). Washington, DC: American Council on Education.

Wittrock, B. (1993). The modern university: The three transformations. In B. Wittrock & S. Rothblatt (Eds.), *The European and American university since 1800* (pp. 303–362). Cambridge: Cambridge University Press.

Wolfe, A. J. (2013). *Competing with the Soviets: Science, technology, and the state in Cold War America*. Baltimore, MD: JHU Press.

Young, M. (1958). *The rise of the meritocracy, 1870–2033*. High Holborn: Thames and Hudson.