

Research Article

A Somewhat Different Perspective on Occupational Prestige Ratings

Ronald Henss¹

1. Independent researcher

This study presents previously unpublished empirical surveys on the assessment of occupational prestige in Germany from the years 1990, 1997, and 1999. The analysis of individual judgment patterns reveals substantial differences with respect to the absolute judgment level and the differentiation among occupations. Nonetheless, an extraordinarily high degree of relative agreement among raters is evident. The average correlation between two raters is .60, the average agreement of individuals with the group judgment is .77, and the reliability of the group standards is nearly perfect. The correlation between the standards of male and female raters, between male and female occupational titles, and between raters from East Germany and West Germany exceeds .96. Comparisons with German surveys from the 1950s to 2018 and studies from China, Great Britain, Poland, Sweden, Spain, South Korea, Hungary, and the USA from 1925 to 2024 demonstrate high to very high levels of agreement. Nevertheless, beneath the surface of these overwhelming commonalities, several systematic temporal trends and local peculiarities can be identified.

Correspondence: papers@team.qeios.com — Qeios will forward to the authors

Introduction

Donald J. Treiman^[1] opens his groundbreaking work „*Occupational Prestige in Comparative Perspective*” with the words: „Men are known by their work. It is no accident that when strangers meet, a standard opening gambit is the question, ‘What sort of work do you do?’ for this information provides the best single cue to the sort of person one is. It marks a person as ‘someone to be reckoned with’ or as someone who can be safely ignored, or to whom deference is due or from whom deference can be expected.”

At all times and in all cultures, the activities that people engaged in were a fundamental determinant of their social status and the associated opportunities for shaping and experiencing life. This is particularly

true for men, who have always performed a more diverse range of activities than women.

An instructive example is surnames in Western cultures. As populations grew rapidly in the Middle Ages, occupational fields became more differentiated, and mobility increased, it was no longer sufficient to identify people by their given names alone. Very often, occupational names were used as an addendum, which were then passed down through generations as family names. Germany is probably the most striking example of this. The 11 most frequently occurring surnames are either occupational titles or derived from professional activities. Of the 50 most common surnames, 33 are occupational names and there are 17 (18) more in the next 50.¹ This means that half of the 100 most common surnames in Germany refer to an occupation.

The significance of one's occupation is not limited to the assignment of family names; it extends to a wide variety of areas of life. Apart from sleep, most men—and increasingly women—spend more time at work than on any other activity during decades of their adult lives, and during their years of schooling, important prerequisites for later occupational activity are established. Private life is also influenced in multifaceted ways by the professional environment and the opportunities and constraints that come with it. For most men, therefore, their occupation is a key determinant of their social status.

Delimitation of the Subject Area

A person's social status is co-determined by countless factors, such as sex, age, ethnicity, family background, place of residence, level of education, income, wealth, physical attractiveness, various character and temperament traits, religious affiliation, interests, and attitudes. The subject of this paper is the social status that results from one's occupation. It deals with the reputation, esteem, recognition, and prestige of occupations in the eyes of the public. In other words, it deals with the subjective evaluation of occupations; it deals with the hierarchical ranking of occupations based on social esteem.

The concept of occupational prestige links the psychological level of subjective evaluation with the sociological level of social ranking. Its psychological perspective distinguishes it from the concept of socioeconomic status (SES), which currently occupies a dominant position in the social sciences and economics. SES originated from efforts to measure occupational prestige using easily ascertainable criteria^[2]. Here, a weighted combination of education level and income has proven to be the best predictor, with education carrying greater weight. Subjectively assessed occupational prestige and socioeconomic status measured using objective parameters are closely related—the correlation is in the range of .80 to .90

—but the two concepts are not identical, and for some questions, one is more informative than the other. This study focuses on the psychological perspective and will not address socioeconomic status.

Another delimitation concerns the method used to measure prestige. The most obvious idea is to present survey participants with a list of occupations and ask them how much they personally esteem the various occupations. This approach is mainly used by polling institutes, but the method often chosen proves to be very problematic. In Germany, for example, the Allensbach Institute for Public Opinion Research has been asking the question „Here are some occupations listed. Could you please select the five that you value most, that you have the most respect for?“ since 1963^[3]. The list comprises about 20 occupations, and the results receive considerable attention in the media and among various interest groups. Figure 1 shows the results of the representative survey conducted in February 2024^[4].



Figure 1. Occupational prestige. Allensbach 2024.

This is clearly not about occupational prestige in our sense of the word. We mean the esteem, respect, appreciation, and reverence associated with occupations that segment society into groups to which one looks up, considers as equals, or looks down upon. Nurses—and especially nursing aides—police officers, and tradespeople are not the groups to which society looks up and treats with particular respect and veneration, and teachers certainly do not enjoy a much higher reputation than university professors. In a 2024 study published on Statista using a different evaluation method, the garbage collector (70%) ranks on the same level as the judge (70%) and above the university professor (64%), while the sewage and wastewater treatment worker (63%) ranks far above the high school teacher (39%). In such studies, occupational prestige in our sense is conflated with the supposed value to society and, above all, participants' self-presentation, in which they seek to portray themselves in a positive light. In particular, the garbage collector and sewage worker reveal that participants consciously seek to „correct“ social inequality in a purportedly just manner. Such studies provide interesting insights into self-presentation and the „woke“ zeitgeist, but the topic of occupational prestige is grossly distorted.² We will not consider such studies.

Since direct questions about personal evaluation yield scarcely usable results, a detour must be taken. The method of choice is to ask participants what esteem various occupations enjoy in the eyes of the public, emphasizing that it is not their own evaluation that matters but public opinion. This concerns opinions about the opinions of others, that is, second-order opinions. This approach largely, though not entirely, circumvents idiosyncratic preferences and aversions as well as self-presentation driven by social desirability.

An important point concerns the precise delimitation of the issue. As already emphasized, occupational prestige involves esteem, reputation, appreciation, respect, reverence, and deference. However, some studies, some of which are highly interesting, focus on other aspects. For instance, a survey in 20 countries inquires about trust in various occupations^[5]. Trustworthiness is certainly related to prestige, but for some occupations there can be striking differences. Another study in 16 countries asks, „Would you be happy or unhappy if one of your children held this kind of job?“^[6]. This question is closer to occupational prestige, but here too, other factors can distort the picture. Apart from the GfK survey, we will not delve into such studies.

Empirical Studies

In the following, we will look at empirical studies on occupational prestige. We will follow the historical development and focus on studies that we will later compare with our own findings.

A hundred years ago, George S. Counts^[7] undertook what is widely regarded as the first empirical survey on occupational prestige. The judges were 82 teachers from Minnesota and 368 students from five different high schools and colleges in Wisconsin and Connecticut. The participants were asked to rank 45 occupations according to their social standing. All pairwise correlations between the group standards³ were greater than .90, the highest being .97; and Counts' conclusion is: „That there are clear-cut differences in the social status of occupations is evident. Whether we like it or not, some occupations, at least according to the views of the groups investigated, receive a much higher social rating than other occupations“ (p. 26).

In the USA, Smith^[8] had 349 students from the classes of 1939 to 1941 rate the prestige of 100 occupations on a 100-point scale using a complex procedure. Two findings are noteworthy for our topic. First, at the top are Supreme Court justices, ambassadors, ministers, senators, and governors—thus, judges and politicians at the highest level.⁴ Second, Smith notes that the occupations at the top and bottom of the list show less variation than those in the middle range, although he does not remark that this is a statistical necessity under regular conditions.

The first milestone in occupational prestige research is a study conducted shortly after World War II at the National Opinion Research Center (NORC) at the University of Chicago under the direction of North and Hatt^[9]. To evaluate 90 occupations, the instruction was: „Please pick out the statement that best gives your own personal opinion of the general standing that such a job has,“ with response options: 1. Excellent standing; 2. Good standing; 3. Average standing; 4. Somewhat below average standing; 5. Poor standing; X. I don't know where to place that one. The study is not spectacular in itself, but it was the impetus for numerous subsequent investigations, and NORC remained the heart of prestige research for decades.

Hodge, Siegel, and Rossi^[10] report a correlation of .99 for a 1963 replication with the 1947 North-Hatt study. In their overview of the state of research, they note that no systematic trends for individual occupations are discernible from 1925 to 1963.⁵ However, there were some minor changes between 1947 and 1963. New scientific occupations emerged and scientific occupations gained prestige, while the esteem for culturally oriented occupations such as concert musicians and radio announcers slightly declined, and craft occupations showed a slight upward trend. The conclusion, however, is: „Nevertheless, the overriding

conclusion must be that the structure of occupational prestige is remarkably stable through time as well as space“ (p. 286).

A comparison across space was undertaken, for example, by Inkeles and Rossi^[11]. Correlations between prestige ratings from the USA, Great Britain, Germany, New Zealand, and Japan ranged from .91 to .97. Even the comparison with the USSR showed remarkably high agreement: Japan .74, Great Britain .83, New Zealand .83, Germany .90, USA .90. Hodge, Treiman, and Rossi^[12] report an average correlation of .91 between the USA and 23 other nations, including several less developed countries from Africa and Asia.

The second milestone and undoubtedly the most significant individual work on occupational prestige is Treiman's^[1] book, the opening words of which we quoted at the beginning. Treiman, who began his research at NORC, first presents a comprehensive theory of occupational prestige. Of paramount importance is his empirical work, in which he developed the Standard International Occupational Prestige Scale (SIOPS) from studies in different regions of the world, which remains a benchmark against which other studies are measured to this day.

The basic theoretical assumption is that differences in occupational esteem inevitably emerge from the division of labor. All human societies exhibit a division of labor among their members, and the more advanced the society, the more diverse and complex the differentiation. Differences in knowledge and specific skills required to practice a profession, differing control over economic resources, and differing authority—that is, differing legitimate control over others' activities—lead to differences in power and privileges, and these differences lead to differing esteem. Since for factual reasons job-specific requirements are almost the same in all societies at a similar stage of development, the hierarchical gradation of prestige is also very similar. In other words, the relative gradation of occupational esteem arises from the nature of the matter, and apart from very simple societies and some local peculiarities, the basic structure is universally very similar.

Treiman's primary achievement is the empirical evidence for the universality assumption. His starting point was 85 studies conducted in the 1950s and 1960s in 60 societies from all parts of the world, including less developed countries from Africa, Asia, and Latin America. Based on the International Standard Classification of Occupations (ISCO) of the International Labor Office, he considered 509 occupations, of which only a very small subset was represented in each study. For countries with multiple studies, he aggregated data if the correlation was at least .95; otherwise, he selected the qualitatively best study. Thus, his data base was a matrix of 60 countries × 509 occupations with a large number of empty cells. Treiman's first key finding concerns agreement within individual societies. „One of the most striking

features of occupational prestige systems in almost all societies is the lack of subgroup variations in prestige ratings. On the average, people in all walks of life, rich and poor, educated and ignorant, urban and rural, male and female, view the prestige hierarchy in the same way. With minor exceptions, there is extraordinary consensus throughout each society regarding the relative prestige of occupations“ (p. 59). Of course, this does not mean that all individuals hold exactly the same opinion. What is important for our topic is the observation that the average correlation between any two individual raters is .60. The second key finding is that the average correlation between any two countries is approximately .80. „Although international agreement is by no means perfect, it is clear that a very strong common dimension underlies the manifest prestige structure of all societies“ (p. 80). After excluding 5 countries in which aspects other than prestige were captured, Treiman constructed a common scale for 509 occupations, which he termed the Standard International Occupational Prestige Scale (SIOPS). The comparison of national scores with the mean of the remaining countries yields an average correlation of .89. This corresponds exactly to the value for West Germany, which is of particular interest to us. Treiman summarizes his central finding as follows: „To sum up, with the possible exception of the least industrialized societies in the world, on the one hand, and the socialist bloc nations, on the other, the evidence just presented suggests that the Standard Scale scores will provide highly accurate estimates of the prestige of specific occupations in any nation of the world“ (p. 179). Two points should be noted in relation to our topic. First, in the communist countries at that time, there was a tendency to devalue office occupations and revalue manual jobs. Secondly, at the top of the SIOPS list, with 78 points, are ‘Physicians, medical and osteopathic’ and the group ‘Teachers, college and university’, which unfortunately is not broken down by subject; at the bottom of the list are ‘Bootblacks’ with 12 points. Treiman explicitly points out that small differences between occupations should not be overinterpreted: „differences of less than six points between the scores for individual occupations should not be regarded as meaningful“ (p. 184). As we will see, this rule of thumb is of limited utility, as statistical significance and effect size also depend on the standard deviation. Also of interest to us is that among 50 frequently considered occupations, the pastor exhibits the largest standard deviation, meaning opinions diverge most widely for him.

Besides Treiman’s SIOPS, the prestige scales of the National Opinion Research Center are of outstanding importance^{[13][14][15][16]}. Based on nationwide surveys in 1963, 1964, and 1965, which were also used to construct the SIOPS, a prestige scale was created for the first time that encompassed all occupations considered in the 1960 National Census^[17]. A methodological innovation of the 1964 survey^[10] was a nine-level „Ladder of Social Standing“ printed on a card, on which participants were asked to arrange cards with occupational titles. The scores were then transformed onto a scale from 0 to 100.⁶

This method was also used a quarter-century later in the 1989 NORC General Social Survey^{[13][13][14]}. Here, 740 occupations were considered. Ten groups rated 110 occupations each, of which 40 were common to all lists and 70 were specific. From the ratings of the 740 occupations, scores were computed for 503 occupational categories of the 1980 Census. One detail of this reduction is particularly interesting to us. The researchers considered university professors so similar in terms of education, income, and activities that they combined them into a single category 'College Professor.' Their accompanying footnote is: „We apparently underestimated the public's awareness of prestige differences among college professors. As an experiment, ten subject-specific postsecondary titles were rated. Their scores were: Professor of Mathematics (78), Professor of Physics in a College or University (75), Professor of Biology in a College or University (74), Professor of Psychology in a College or University (74), Professor of History (73), Professor of English (72), Professor of Business Administration in a College or University (71), Professor of Foreign Languages (70), Professor of Social Work in a College or University (66), and Professor of Drama (62)“^[14]. The obvious fact that there are considerable differences in the reputation of professors depending on the subject is ignored in nearly all studies. From the multitude of findings, only a few that are relevant to our topic will be mentioned. First, for the 40 occupations rated by all, the average correlation between any two raters is .40. Of the 1,166 participants, 41 show a negative correlation with the group standard (-.006 to -.975). After their exclusion, concordance is .45. Second, for 160 occupations also rated in 1964, the correlation with the Nakao-Treas prestige scale is .97. Third, despite the extraordinary relative stability over a quarter-century, 71 of the 160 occupations show a significant change in absolute rating, with 57 improving and 14 declining. Fourth, the minimum rose from 13.7 to 19.1, and it is primarily lower-status occupations that increased their esteem, although they remain at the bottom. „[L]ower-status occupational titles did gain disproportionately from the upgrading. The American public no longer viewed workers at the bottom of the occupational ladder as being so distant from those on the middle rungs“^[14].

An update using the same method was conducted as part of the 2012 General Social Survey with 1,001 participants^[18]. Now, 860 occupations were considered, divided into 12 sets of 90 occupations, of which 20 were common to all sets. The number of raters per set ranges from 74 to 96.

Surprisingly, particularly rich information on the temporal stability of occupational prestige comes from Poland. Pohoski, Słomczyński, and Wesołowski^[19] report on 11 studies from 1958 to 1973 with large urban and rural samples, and a nationwide survey from 1975. Between 9 and 42 occupations were rated on a 5-point scale. Correlations between group standards range from .88 to .97, averaging .94. Domański, Sawiński, and Słomczyński^[20] added 8 nationwide surveys, one from 1987 (still in the socialist phase) and

7 from the post-socialist period 1991 to 2008. The correlations (1958/1987) and (1975/1987) are .88 and .95. Thus, the socialist phase is characterized by a very high degree of stability overall. A comparison of the 7 group standards from the post-socialist period yields values from .61 to .96, averaging .81. This period is thus characterized by lower stability. However, this applies primarily to the first survey after the upheaval in 1991, which correlates with the others on average only .73; for the other time points, the average correlation is .82 to .88. The lower stability is largely due to radical shifts in political occupations. „The minister has moved from the top to the bottom of the hierarchy of prestige among occupations and has been reduced to the level of an employee, an unskilled worker, or a cleaner. This degradation extends to all representatives of the ruling class“^[20].⁷ The pastor has also suffered a sharp decline in prestige. Two aspects that are rarely considered in occupational prestige research are of particular interest to us. The average correlation between pairs of raters in the years 1967, 1987, 1996, 2004, 2006, and 2008 was .38, .32, .23, .18, .22, and .21, respectively. In many areas involving subjective judgments, these would be high values, but for occupational prestige ratings, as Treiman^[1] already noted, the average interrater correlation is about .60. The exceptionally low agreement is certainly partly due to scale usage. In 1958, 21.0 percent of ratings fell into the highest category of the 5-point scale, but by 2008 this had fallen to only 12.4 percent. For the lowest category, the percentage fell from 5.5 to 2.9. In 1967, 55.0 percent of respondents used the full scale width from 1 to 5; in 2008, only 18.8 percent. In 1967, 0.5 percent used only two adjacent categories, compared to 11.6 percent in 2008. „More and more respondents do not differentiate occupations by prestige at all, increasingly restrict the rating scale to a small range, while fewer and fewer are willing to differentiate prestige to the extent the scale allows“^[20].

A Chinese prestige hierarchy from 1993/1994 is provided by Chiu^[21]. Participants were 1,196 adults from four urban districts in Guangzhou. A total of 102 occupations were considered, but participants rated only one of three sets on a scale of 1 to 6 and were then asked to give reasons for their rating for one of the occupations. Bian^[22] had shown that the prestige hierarchies from Beijing in 1983, Tianjin in 1988, and Taiwan in 1988 correlate with Treiman's SIOPS and the NORC at a level of .85 to .89. We will limit ourselves to Chiu's data, as her study covers more than twice as many occupations.

Garcia-Mainar and Montuenga^[23] report on two surveys conducted in Spain in 1991 and 2013. However, they only share the results at the level of 66 main groups and 207 occupational groups. For the main groups, the 1991/2013 comparison yields a correlation of only .734, and at the occupational group level, only .556. Of the 207 occupational groups, 21 improved by more than 50 ranks and 24 declined by more than 50. The most spectacular changes concern politicians. Members of regional governments have plummeted

from rank 3 to the penultimate position at 206, while legislators and senior government officials have fallen from rank 5 to the very last place. Also worth mentioning are religious occupations, which fell by almost 100 places, and, on the other hand, occupations in the care sector, which improved by just over 100 places, and firefighters, who rose by 73. For our own analyses, we rely on data from CIS^[24] for 2013. Nineteen groups of 314 participants each rated 15 occupations on a scale of 0 to 100. That makes a total of 285 occupations.

In representative surveys in 2002 and 2018, 1,819 and 1,653 participants in Sweden, respectively, rated the social status of 100 and 120 occupations („how it is valued in society in terms of status“) on a 9-point scale. For 91 occupations, the correlation between the scales 16 years apart is .974, and the correlation with Treiman's SIOPS is .826 and .819^[25]. Primarily the lower occupations gained esteem; the minimum rose from 8 to 14. Thus, the gap between top and bottom was narrowed without altering the ranking.

Hughes, Srivastava, Leszko, and Condon^[26] used the US Department of Labor's O*NET database to determine the prestige of 1,029 occupations. The survey was conducted online in 2015/2016 and included more than 3,000 participants, who rated an average of 52 occupations. The ratings were given on a scale of 0 to 100 using a slider. On average, each occupation was rated by 150 participants. The average correlation between groups that had rated the same occupations is .93. The correlation with NORC 1989 and NORC 2012 is .79 and .85, which is the same order of magnitude as the correlation between the two NORC scales, which is .80.

As part of the 2016 Hungarian Microcensus, a total of 173 occupations were considered. 7,008 participants ranked 15 occupations each, and values were transformed to a 0 to 100 scale based on relative positions. Besides prestige, occupations were also ranked by knowledge, power, income, usefulness, and zeitgeist.⁸ Huszár, Hajdu, Sik, and Nagy^[27] report results at the occupational group level. Own calculations at the level of the 173 occupational titles using data from Sik, Huszár, and Nagy^[28] show a correlation of .95 for prestige with knowledge, .91 with power, .81 with zeitgeist, .78 with income, and .62 with usefulness to society. At the occupational group level, the picture is the same, with slightly higher coefficients. We will only consider the level of job titles and only occupational prestige.

In South Korea, in an online survey in 2022, 2,132 participants rated the prestige of 32 occupations on a 0 to 10 scale^[29]. Correlations between subgroups separated by university degree, leadership position, and high household income (yes/no in each case), are all greater than .998. At the top is the politician (Congressperson, mean 8.18), closely followed by the lawyer (8.05) and the physician (8.04). Particularly noteworthy is that the politician also has the largest standard deviation. We will return to the relationship

between mean and standard deviation in detail later. For 15 occupations also rated in 1990, 2000, 2009, and 2016, rank correlations between survey time points range from .925 to .993. Thus, the prestige hierarchy in South Korea also shows extraordinary stability over three decades.

In Great Britain, Newlands and Lutz^[30] examined the prestige and social value of 576 occupations, including several with very low esteem (e.g., street sweeper, garbage collector) and some illegal (e.g., drug dealer, human trafficker) plus some not belonging to the regular labor market: unemployed, retiree, student, and homemaker. The data collection was conducted via the Internet in March 2022. The ratings were made using a slider on a scale from 0 to 100. 1,219 participants rated prestige and 1,210 rated social value. Participants were left to their own understanding of prestige and social value. Each participant rated 245 or 246 occupations,⁹ and on average, each occupation was rated by 152 participants. The correlation between occupational prestige and social value is .81. Thus, the two concepts are closely related but clearly distinguishable. Gmyrek, Lutz, and Newlands^[31] extended the perspective to artificial intelligence. In 50 runs, GPT-4 was to rate occupations on a 0 to 100 scale, as a sample representative of the social and demographic heterogeneity of the United Kingdom would do. The correlation with human assessments is .92 for prestige and .89 for social value. For a core list of 130 occupations, AI prestige ratings show distinctly higher agreement with Whites than with non-Whites and, within Whites, slightly higher agreement with women than with men (.94 vs. .88).

Artificial intelligence from the GPT-3 family¹⁰ was also used by de Vries, Hill, and Ruis^[32]. Here, 99 occupations were presented in complete pairwise comparisons, and the program was to indicate which occupation performed better each time. For the status criterion, the prompt was:

There are two occupations:

- 1) [OCCUPATION 1].
- 2) [OCCUPATION 2].

Answer only 1 or 2.

The highest status is generally perceived to be: Answer: [1/2]

The same pattern was used for:

The most prestigious is generally perceived to be:

The one who would generally defer to the other would be:

The one Alex is most likely to boast about making friends with is:

The job that performs best on the criteria of pay and how much education or training it requires is:

To control for order effects, there were four runs, with the order of occupations and answers reversed for each pair. In total, there were 48,510 queries. Scores were determined based on the frequency of preference over other occupations. Responses to the third criterion—deference—proved unusable. Status and prestige were intended as alternative indicators of occupational prestige, and the correlation between scores is .99. The correlation with boast-worthiness is .78 and .75, and with training/pay .74 and .72, and the correlation between these two is .49. Of interest to our topic is the correlation of the combined status/prestige scale with Treiman's SIOPS, which is .85. As with Gmyrek, Lutz, and Newlands^[31], this is impressive evidence that artificial intelligence prestige ratings are just as good as those made by human judges.

Since we conducted surveys in Germany ourselves, we now turn our attention to Germany. Here, occupational prestige research is only weakly developed. Essentially, there are only two significant studies. Before considering these, we would like mention two studies from the 1950s.

The first is found in Bolte's^[33] book, which provides insight into the national and international state of research up to the 1950s and whose data also entered Treiman's SIOPS. Bolte surveyed 1,529 individuals from Schleswig-Holstein and 83 from Hamburg. The urban samples included three groups of vocational students—retail clerks, precision mechanics, bank and insurance clerks—and a group of economics students. These were to rank 38 occupations. In a rural sample from three remote villages, self-employed farmers, self-employed craftsmen, farm workers, and rural teachers were surveyed. With these, extensive interviews were conducted, in which interviewers ranked 47 occupations. Bolte's conclusion is: „In principle, the scales in city and country agree“ (p. 59). We can fully confirm this. Our own analysis yields correlations between urban samples of .96 and .98, between rural of .94 to .99, and between the two overall groups .96. It should be noted that opinions regarding pastors vary widely.

Kleining and Moore^[34] captured the prestige of 70 occupations in the late 1950s. 150 raters aged 16 to 65 „were presented with a six-level ladder on which they had to place the individual occupations written on cards ‘according to their esteem from very bottom to very top’“ (p. 519). There are no differences in the relative status of the occupations between male and female assessors, nor between low- and high-income workers, nor between low- and high-income civil servants, employees, and the self-employed. In absolute rating, there are a few minor differences; for example, low-income workers tend „to rate all occupations somewhat lower than the other respondents, especially higher occupations, while lower ones, to which they assign themselves, are rated in agreement with the others,“ and „middle-class occupations and those from non-industrial sectors are generally rated somewhat higher in the city. In contrast, industrial

occupations and lower-class job are rated somewhat higher in rural areas than in the city. The differences are small but noticeable“ (p. 551). Treiman^[1] reports a correlation of .956 with SIOPS.

The only significant German scales known to us are, on the one hand, the Magnitude Prestige Scale (MPS) by Bernd Wegener^{[35][36]} and, on the other hand, the Berufliche Ansehensskala (BAS; Occupational Esteem Scale), which was developed more recently^{[37][38]}.

Wegener's Magnitude Prestige Scale differs methodologically from all others. In two representative studies in 1979 and 1980, 4,015 participants assessed 50 occupations on „how much people in these occupations are esteemed in our society today“^[35]. The assessments were made both on a 9-point rating scale and according to the magnitude method, which was developed in psychophysics^[39]. Here, one occupation serves as the standard stimulus, and the participant assesses all others relative to this standard.¹² Thus, in contrast to category scales, the scale is open-ended upward. We will return to this peculiarity later. Based on the relation between categorical scores and magnitude scores, as well as the relation to corresponding scores from Treiman's SIOPS, prestige was computed for additional occupations that had not been assessed at all, resulting in a total of MPS scores for 283 occupations. For these, the correlation between MPS and SIOPS is .858. For our later considerations, it should be noted that Treiman's SIOPS was used in constructing the MPS scores.

The most comprehensive German study on occupational prestige was conducted in the research project „Berufe in Deutschland: Gesellschaftliche Wahrnehmung und Persönlichkeitseigenschaften“ (Occupations in Germany: Societal Perception and Personality Traits) of the Bundesinstitut für Berufsbildung (Federal Institute for Vocational Education and Training) under the direction of Christian Ebner and Daniela Rohrbach-Schmidt^[38]. Here, 402 occupations and a representative sample of 9,011 participants were considered. The survey was conducted from October 2017 to May 2018 as computer-assisted telephone interviews. Each participant rated only 5 occupations, as other characteristics were also considered besides prestige.¹³ Due to the very large sample, each occupation has 67 to 145 raters. The following findings are of interest for our topic. First, the comparison of the first and second survey phases yields a test-retest reliability of .93. Second, the highest prestige value is 8.65 (chief physicians) and the lowest 3.12 (telephone operators). This results in a range of 5.53. Since the scale runs from 0 to 10, 44.7 percent of the range is unused. Third, the correlation with Wegener's Magnitude Prestige Scale is .62, and the correlation with 258 occupations from Treiman's SIOPS is .72. Fourth, after master butchers (2.94) and illustrators (2.84), pastors have the third-largest standard deviation at 2.78 and rank 360 among 402 occupations.

Firefighters rank 7th. General practitioners rank 13th, but physicians only 38th. Midwives and obstetricians rank 18th, but nurses rank 127th.

In conclusion, we summarize some points relevant to our topic. First, adults and adolescents have no difficulty assessing occupational prestige. This indicates that they possess established standards that they can retrieve on request. Second, agreement among individual raters is very high. This necessarily implies that the vast majority of individuals also agree highly with the group judgment, and a few dozen raters suffice to obtain an extraordinarily reliable group standard. Third, different groups, regardless of composition, show extraordinarily high agreement in the relative evaluation of occupations. This applies particularly also to the comparison of different nations. Fourth, the occupational hierarchy is extraordinarily stable over time. Fifth, despite temporal stability, there are notable shifts and systematic trends for individual occupations and occupational groups, but these do not entail fundamental changes in the hierarchy. Sixth, one such trend is the improvement in esteem for the lowest occupations, even if they remain at the bottom.

As benchmarks, we note four figures from Treiman's groundbreaking work. First, the average correlation between two individual raters is approximately .60. Second, the average correlation between occupational prestige scales within the same country is often around or above .95. Third, the average correlation between any two countries is approximately .80. Fourth, the average correlation between the SIOPS and prestige scales of other countries is approximately .90. Regarding the last two benchmarks, it should be noted that Treiman considered countries at different stages of development from all continents and that a strong communist bloc still existed at that time.

Objectives of This Study

At the end of the 1980s, we began investigations of physical attractiveness ratings based on facial photographs^{[40][41]}. Building on this, we developed a theory of mate value grounded in evolutionary psychology and extended face research to the topic of face and personality impressions, incorporating factorial models from differential psychology^{[42][43][44][45]}. Since social status is a central, if not the most important, component of men's mate value, in 1990 we conducted a study in which participants were to rate the prestige of 95 occupations. This survey was replicated in 1999 in an online study. Additionally, in 1997 as part of an experiment on mate selection criteria, we collected prestige ratings for 40 male and 40 female occupations. All these surveys were used solely for internal purposes, and results have not been published to date.

As can be seen from the brief description, occupational prestige is not one of our areas of expertise; and we do not want to „sell“ a new kind of prestige scale. Given the very high temporal and spatial stability of the prestige hierarchy, such a thing is not possible. This study has two main objectives. Our surveys were conducted between 1990 and 1999. Thus, they lie between the data collections for the Magnitude Prestige Scale (MPS) 1979/1980 and the Berufliche Ansehensskala (BAS) 2017/8. Including the two studies from the 1950s opens a comparison spanning a period of dramatic political and societal change in Germany and radical transformation in occupational structure worldwide. We will extend the comparison to the international studies mentioned above. Our second main objective is methodological. We aim to examine more closely aspects that are only marginally or not at all addressed in the literature. This concerns, on the one hand, individual judgment patterns and a differentiated consideration of rater agreement in prestige ratings, illuminating consensus at different levels. On the other hand, we aim to focus more than usual on comparisons between individual occupations. Among other aspects, we want to draw attention to the fact that absolute and relative comparisons can lead to surprising divergences in individual cases and that differences between not-too-distant occupations should be interpreted with caution.

Methods

In this work, we report on three of our own studies. The first was conducted in 1990, the second in 1997, and the third in 1999. Our primary focus is the 1999 survey, in which only male occupations were rated.

Experimental Procedure

First, we consider the main study conducted at the end of 1999, thus capturing occupational prestige at the end of the 2nd millennium. This survey was conducted as an online questionnaire, and participants arrived via self-selection through links on our homepage at Saarland University. Participation was voluntary, anonymous, unpaid, and non-binding.

Under the heading „Psychological Online Survey: Esteem of Occupations. Saarland University. Department of Psychology,“ the following instruction was found.

In the following, you see a list of occupations. Please indicate for each occupation what **esteem it possesses in the eyes of the public**, specifically with reference to men. Assign a **number from 0 to 100** to each occupation. Here, 0 stands for the occupation with the very lowest esteem and 100 for the occupation with the very highest esteem. Accordingly, for example, 25 denotes fairly low esteem, 50 medium esteem, and 75 fairly high esteem. Please

also use the intermediate numbers. If an occupation is completely unknown to you, please enter -1. Please note: It does not matter whether you personally would like to perform the occupation in question—what matters is the esteem the occupation enjoys **in public opinion**.

Below was a table in whose left and right halves 49 occupations each were listed vertically with a field for entering the estimated prestige. The order was constant throughout the survey. The selection of occupations was intended to represent a very broad spectrum. The 0 to 100 scale was deliberately chosen so that ratings could be associated with percentages.

Following the ratings, participants were to indicate their sex, age, origin, occupation, and the prestige of their own occupation. For origin, a list of German federal states and Austria, Switzerland, Luxembourg, and Other was provided.

We had already conducted an analogous study in 1990. At that time, before the internet era, data were collected in face-to-face interviews by students at Saarland University. The list comprised 95 occupations, all of which were also included in 1999. Additionally included in 1999 were ambassador, business administration student, and educational counseling director. A list of the 98 German occupational titles with their English counterparts is provided in the appendix in Table A5. Unlike in 1999, in 1990 the occupations were rated in two parallel versions, one in the male and one in the female form. For brevity, we will speak of male occupations and female occupations and explicitly point out that this does not refer to male-typical and female-typical occupations in terms of relative frequency. Due to multiple changes in computer systems, operating systems, and storage media, the raw data are no longer available today. The only data preserved are the means.

The 1997 occupational prestige ratings are a byproduct of an experimental study on multiple criteria of mate choice. The computer-controlled experiment was conducted in individual sessions under the supervision of a female experimenter at the Psychological Institute of Saarland University.¹⁴ Following the actual experimental manipulation, participants were to rate, among other characteristics, the prestige of the 40 male or 40 female occupations used in the experiment. The instruction followed the pattern of the other two studies. In accordance with the research question, male occupations were rated only by women and female occupations only by men. A detailed evaluation of the two sub-experiments is found in the diploma theses by Rita Breinig^[46] and Isabel Schmidt^[47]. For the present work, we reanalyzed the occupational prestige ratings independently.

Participants and Dataset

The 1990 survey was conducted with students from various disciplines at Saarland University. Since the raw data are no longer available, we cannot provide further details. We can, however, note that the means on which we rely have extraordinarily high reliability. This will become immediately evident in the analysis of rater agreement in the other studies.

The two 1997 experiments on multiple criteria of mate choice involved 75 male and 73 female students from Saarland University. Male ages ranged from 22 to 33 years, with a mean of 25.1. Females, with a range of 19 to 27 years and a mean of 23.3, were somewhat younger. In both samples, there were 29 psychology students; otherwise, the fields of study are almost disjoint. The vast majority of women studied humanities, while the men predominantly studied STEM fields.

In the 1999 online survey, after a critical check, in which a few obviously unusable submissions were removed, a dataset of 136 participants remained. Table 1 shows a breakdown by origin and sex.

Origin	Males	Females	Total
West Germany	45	59	104
East Germany	7	6	13
Austria	6	9	15
Luxembourg	1	1	2
Other	1	1	2
Total	60	76	136

Table 1. Sex and origin of participants.

104 participants are from western federal states, only 13 from eastern, 15 from Austria, 2 from Luxembourg, and 2 from other countries. With 76 to 60, the number of women is slightly larger than that of men. With a mean age of 28.8 and an interquartile range of 23.0 to 33.3, men are 3.7 years older than women (mean 25.1, interquartile 20.0 to 28.3). 60 students make up the largest groups; otherwise, participants are spread across various occupations, most of them in academia or with an academic

qualification as a prerequisite. Thus, the sample reflects the typical pattern in online studies with self-selected participants, especially at that time when internet access in non-academic environments was much more limited than today.

Results

In the data analysis, we first focus on the raters. We examine, on the one hand, the response styles of individual raters and, on the other, consensus among raters.

Intraindividual Perspective

We begin with the individual raters. For each participant, the mean (M) and standard deviation (s) across all occupations were computed. Table 2 shows statistical descriptors of these two parameters. PR25, PR50, and PR75 denote the 25th, 50th, and 75th percentiles.

	M	s
Mean	52.1	23.5
Standard deviation	6.39	4.81
Minimum	33.5	14.4
PR25	47.3	20.2
PR50	52.1	23.5
PR75	56.6	27.0
Maximum	65.3	34.5

Table 2. Statistical descriptors of individual rater parameters.

The mean of the individual means (which is also the mean across all occupations) is 52.1, with a standard deviation of 6.39. The „most pessimistic“ participant accorded occupations an average prestige of only 33.5, while the „most optimistic“ saw average at 65.3. The middle 50 percent range from 47.3 to 56.6.

The standard deviation of individual raters is a measure of how much they differentiated between the various occupations. On average, it is 23.5. The minimum is 14.4, the maximum 34.5. The interquartile ranges from 20.2 to 27.0.

Mean and standard deviation are not independent of each other. The correlation is $-.253$ ($p = .003$), meaning that participants giving higher average ratings show slightly less differentiation.

Overall, we observe enormous differences among raters both in average rating and in differentiation between occupations.

For a detailed examination, we formed three groups of 34 raters each based on the mean: the lower quarter, middle quarter, and upper quarter. Within each group, we determined the distribution of individual judgments. The smoothed frequency distribution of the 34 (raters) $\times$ 98 (occupations) scores is depicted in Figure 2.

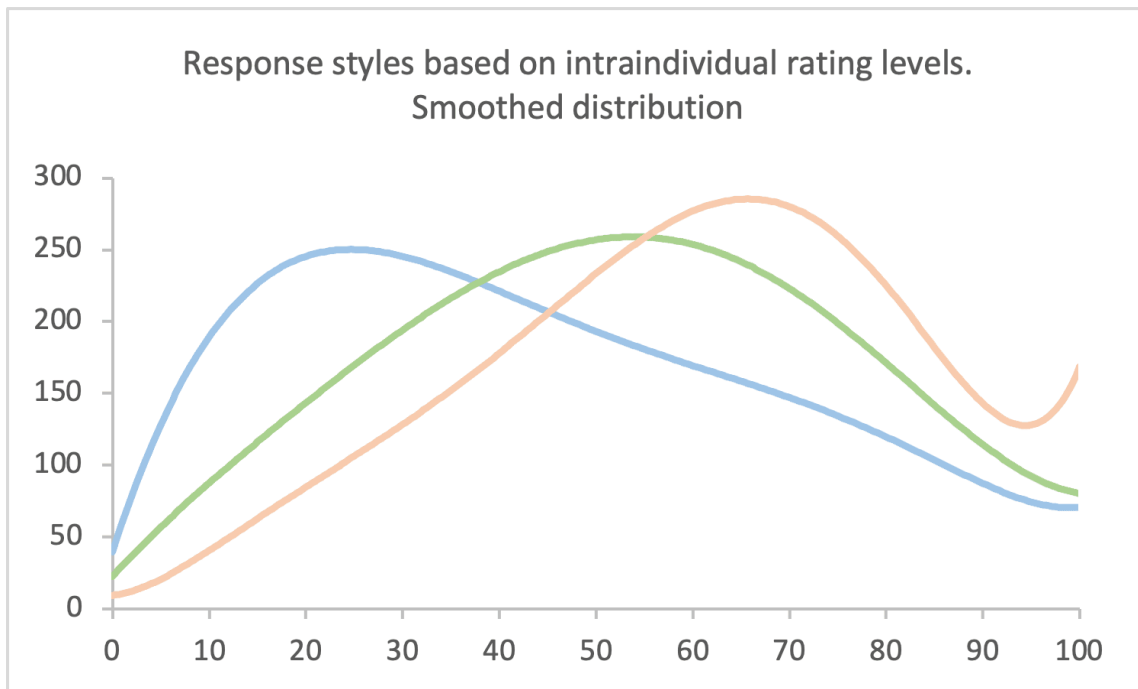


Figure 2. Response styles based on intraindividual rating levels.

The group with the lowest ratings shows a left-skewed distribution with a uniform, relatively flat decline toward the upper end. The middle group shows a rainbow-shaped distribution. The upper group shows a weakly right-skewed distribution with a marked preference for the maximum 100, assigned in 5.6 percent of cases. In the lower and middle groups, the corresponding percentages are 2.4 and 2.8. The lower and

middle groups are also very similar at the minimum score of 0, with 1.1 and 0.9 percent, whereas the upper group is far below at 0.2 percent. The average ratings are 44.0, 52.4, and 60.0.

Analogous to the analysis of means, we formed three groups based on intraindividual standard deviations. The result is shown in Figure 3. Note that the colors denote different groups than in Figure 2.

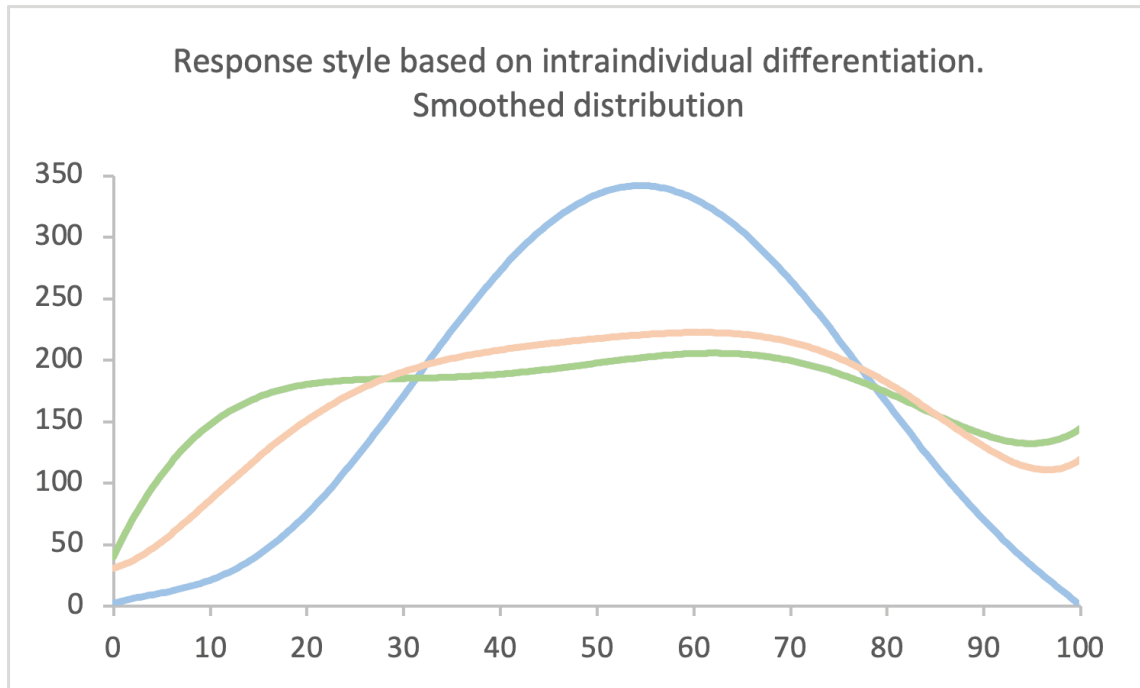


Figure 3. Response style based on intraindividual differentiation.

The group with the least differentiation (blue) shows a bell-shaped distribution with a vanishingly small proportion of extreme ratings. The middle group (orange) shows a flattened distribution that declines above 80 and below 30. Particularly interesting is the group with the strongest differentiation (green). It shows a flat distribution that covers almost the entire spectrum relatively evenly and drops sharply only in the lowest range. The intraindividual standard deviations average 17.5, 23.5, and 29.7.

For the question of the occupational prestige hierarchy, these groups are of very different relevance. Levelers contribute little to differentiation. In comparison, the influence of the middle group is considerably greater. Even greater is the influence of differentiators; their judgment carries a disproportionate weight due to the very broad dispersion.

The enormous differences in mean-level and differentiation are not unique to occupational prestige ratings; they occur everywhere in subjective evaluations, but this perspective is rarely taken into account.

Consensus Among Raters

Interest typically lies not in individual participants but in the reliability of the group standard derived from the average across all raters. To this end, we now examine consensus among raters with respect to the relative evaluation of occupations.

In a reliability analysis, first, the average correlation between any two raters ($r(i,j)$) was computed. Second, the correlation between each rater and the group standard of the remaining participants ($r(i,rest)$), in the language of test theory, this is the discrimination of individual raters. Third, internal consistency, i.e., reliability of the group standard (Cronbach's alpha and MacDonald's omega). These analyses were conducted for the 1999 main study, both for the total group and for the male and female subsamples. Additionally, the correlation between group standards was calculated ($r(M,F)$). In the 1999 study, this is the comparison of male and female raters; in the 1997 study, it is the comparison of male and female occupations rated by men and women, respectively. The results can be seen in Table 3. N denotes the number of raters.

Year	Incumbent	Rater	$r(i,j)$	$r(i,rest)$	alpha	omega	$r(M,F)$	N
1999	Male	Total	.594	.779	.995	.995		136
1999	Male	Male	.564	.758	.987	.988	.972	60
1999	Male	Female	.627	.794	.992	.993		76
1997	Male	Female	.710	.840	.994	.995	.970	73
1997	Female	Male	.656	.807	.992	.993		75

Table 3. Interindividual concordance, discrimination, internal consistency, correlation between group standards, sample size.

The average correlation between any two individual raters in the 1999 main study is .594. Thus, 59.4 percent of the variance is true variance attributable to consensually perceived differences in the prestige of various occupations.¹⁵ Figure 4 shows the smoothed frequency distribution of concordance values.

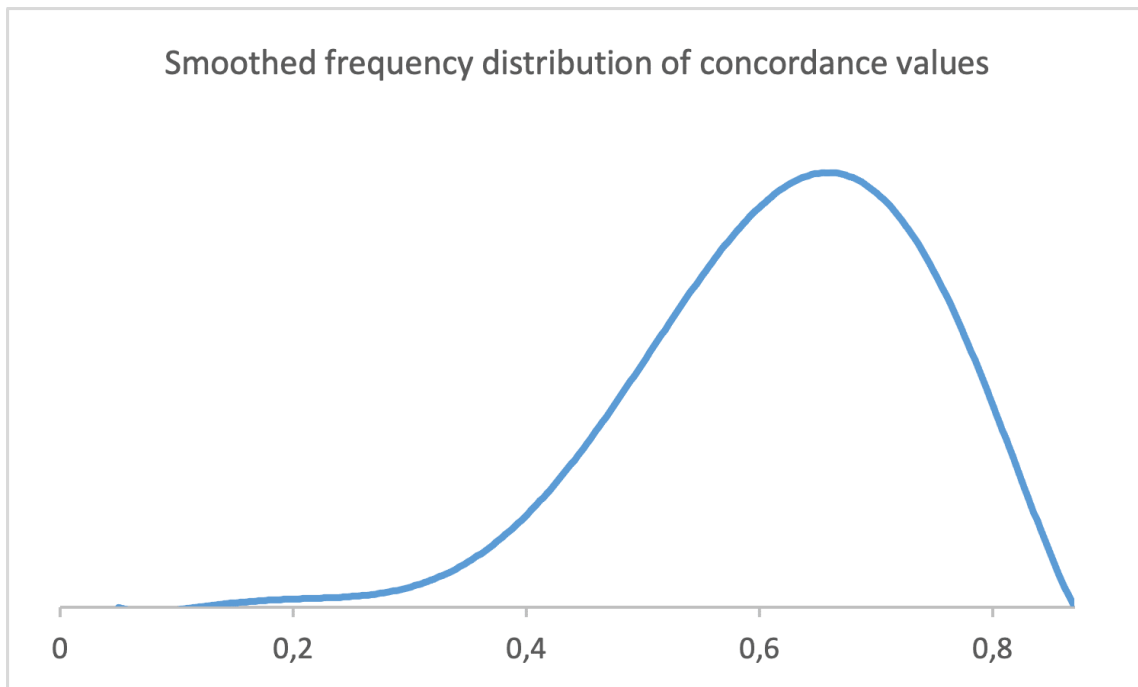


Figure 4. Smoothed frequency distribution of concordance values.

The values are unanimously positive, the minimum is .134. Only 1.4 percent of values are less than .30, and only 7.0 percent less than .40. 52.6 percent of values are greater than .60, and the maximum is .879. Agreement is slightly higher among women than among men (.627 vs. .564), but even among men, the proportion of true variance is nearly sixty percent. In the two sub-experiments in 1997, the concordance was even higher. For male occupations rated by women, it is .710; for female occupations rated by men, .656. In both studies, consensus is slightly higher among women, but for the mate choice experiments, note that only 34 of the 40 occupations were common, so the comparability is limited.

Individual agreement with the average judgment of the other participants is necessarily much higher than interindividual concordance. The group standard is shared to a high degree by nearly all individuals. Across the various surveys and subgroups, values average between .758 and .840.

Due to the extraordinarily high agreement already at the individual level, the group standard exhibits nearly perfect reliability. If the study were repeated with an equally large sample from the same population, one would expect a correlation of .987 with the present group standard in the worst-case scenario, that is, for the group of 60 men from the main study.

The group standards of male and female raters in the main study show a nearly perfect correlation of .972. The same agreement (.970) holds for the comparison of male and female occupations rated by the opposite sex in 1997.

Finally, we revisit the three groups each that differ massively in intraindividual mean or standard deviation. The correlation between their group standards is .974, .967, and .984 for means and .983, .986, and .977 for standard deviations (respectively group 1 vs. group 2, group 1 vs. group 3, group 2 vs. group 3). Thus, nearly perfect agreement is evident here as well.

The overall conclusion is: In absolute rating, there are profound differences among individual raters, both in level and differentiation. With respect to the relative position of occupations, however, there is extraordinarily high agreement among raters, and group standards possess virtually perfect reliability.

Focus on Occupations

In our previous considerations, raters were in the foreground. Now we focus on occupations. We initially limit ourselves to the 1999 main study. The 1990 and 1997 surveys will be considered later in connection with specific questions.

Participants had the option to mark occupations whose prestige they could not assess with -1. This occurred three times each for ministerial councillor and druggist, twice for dispatch manager, and once each for payroll accountant, social pedagogue, and educational counseling director. Otherwise, there were no missing values. The missing values were replaced via regression based on the group standard.

Table A1 in the appendix contains a complete list of occupations with mean, standard deviation, 95% confidence interval of the mean, and skewness. Table 4 summarizes statistical descriptors across all occupations. We limit ourselves to mean and standard deviation, which are of particular interest.

	M	s
Mean	52.1	16.6
Standard deviation	18.1	3.1
Minimum	17.7	8.5
PR 25	40.3	14.7
PR 50	50.5	16.1
PR 75	65.6	18.2
Maximum	91.7	28.6

Table 4. Statistical descriptors of mean and standard deviation across all occupations.

Overall, prestige values range from 17.7 to 91.7, and the middle 50 percent from 40.3 to 65.6; thus, we have a broad distribution. The median is 50.5, the arithmetic mean 52.1.

The standard deviation, i.e., the variability of ratings by different participants, varies greatly across occupations. It ranges from 8.5 to 28.6. The middle 50 percent exhibit a standard deviation from 14.7 to 18.2.

A Key to Understanding Prestige Ratings

After the overall overview, we now look at Figure 5, which shows the mean and standard deviation of individual occupations, marked by dots.

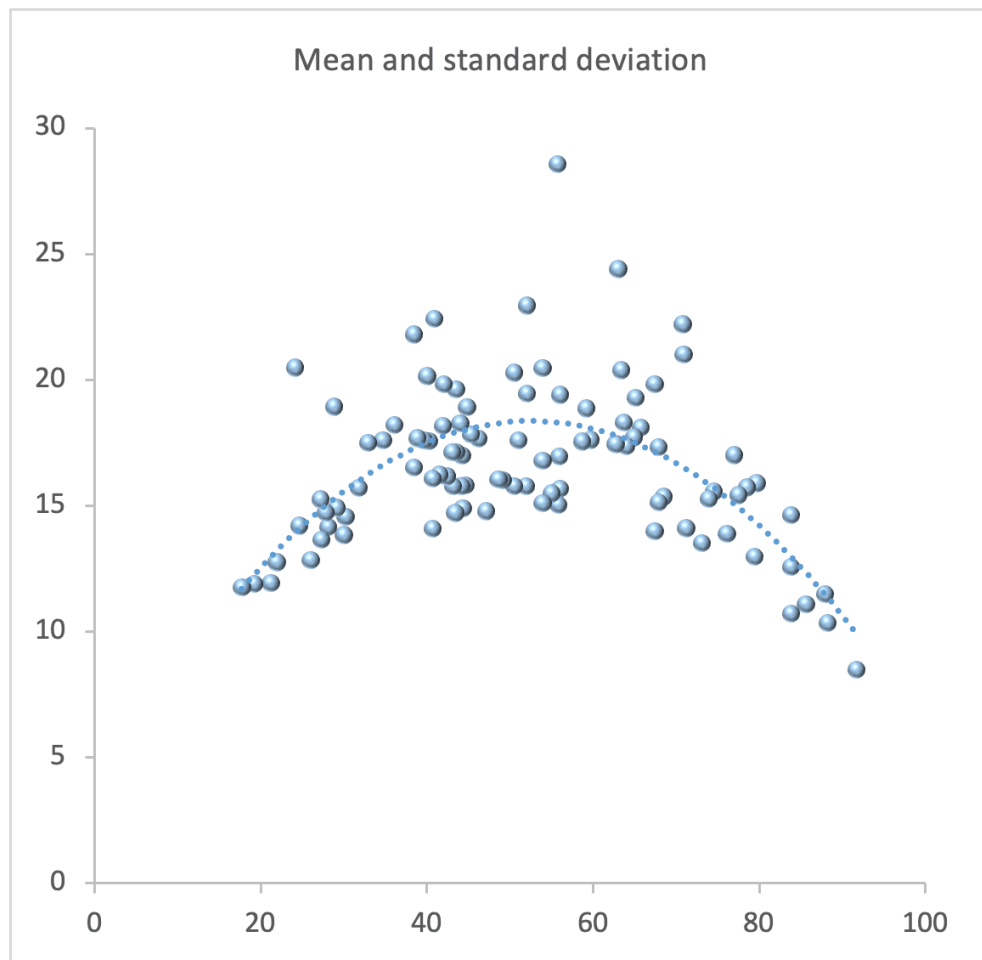


Figure 5. Mean and standard deviation.

The x-axis spans the full scale width. This makes clear that the best-rated occupations are near the maximum, while the worst-rated are well above the minimum. The distance to the extremes is more than twice as large at the bottom as at the top. The central point is the shape of the relationship, described by the dotted line. This denotes the quadratic regression of the standard deviation on the mean. The correlation R is .636, explained variance 40.4 percent. The inverted U-shape shows that dispersion is much smaller at the ends than in the middle. This corresponds to the usual pattern arising from the nature of the matter. One can only be very close to the upper or lower end if the dispersion is very small; however, in the broad middle field, ratings can be very uniform or widely spread. This relationship must be considered in interpreting prestige data. We will illustrate this with concrete examples.

Occupational Prestige: Top, Middle, Bottom

Next, we want to get a rough overview. To this end, we examine the top, middle, and bottom 10 in Figure 6. The complete list will be examined in the next section. A primary goal of the current and following sections is to show that differences between occupations must be interpreted with caution.

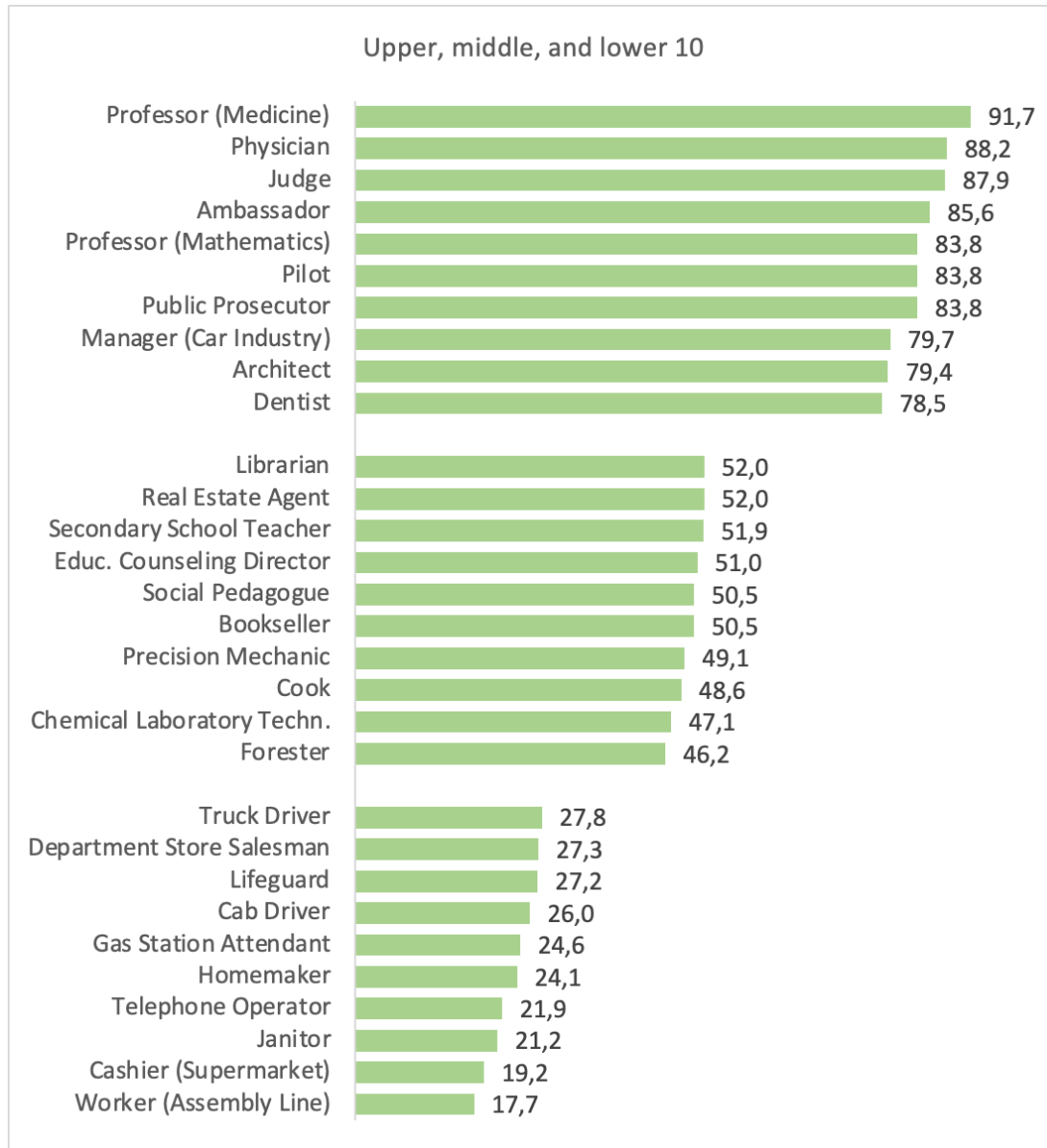


Figure 6. Upper, middle and lower 10.

In this representation, two points are immediately apparent. First, there are vast differences between the three blocks. Second, within the blocks, the differences are relatively small, particularly in the middle

block.

At the top, the professor of medicine leads with a value of 91.7, followed by the physician with 88.2. The rest of the top ten range from the judge (87.9) to the ambassador (85.6), pilot (83.8), public prosecutor (83.8), professor of mathematics (83.8), manager in the car industry (79.7), architect (79.4), and dentist (78.5). The difference between the first and tenth positions is 13.2 points. Statistically, the 3.5-point difference between the first and second positions, the professor of medicine and the physician, is significant at this sample size ($p < .001$). The effect size, measured with Cohen's d , is 0.42. Thus, the gap between the gold and silver medals represents a medium effect. As we will see shortly, this is an exception. To get a sense of the statistical differences, we consider the differences between the three types of professors: the medical professor in first place (91.7), the mathematics professor in seventh place (83.8), and the art history professor in twelfth place (77.0). All pairwise comparisons are significant at the 0.1 percent level. The effect size d is 0.46 (mathematics vs. art history), 0.67 (medicine vs. mathematics), and 0.93 (medicine vs. art history). Thus, we find medium to very strong effects here.

Now we turn our attention to the lower end. The bottom position is held by the worker on the assembly line (17.7), followed by the cashier in the supermarket (19.2), janitor (21.2), telephone operator (21.9), homemaker (24.1), gas station attendant (24.6), cab driver (26.0), lifeguard (27.2), department store salesman (27.3), and truck driver (27.8). Here, the range is 10.1 points. For three pairwise comparisons: truck driver vs. worker on the assembly line ($p < .001$; $d = 0.70$), truck driver vs. gas station attendant ($p < .001$; $d = 0.56$), gas station attendant vs. worker on the assembly line ($p = .013$; $d = 0.22$). Thus, the differences in the bottom ten are not quite as large as in the top ten.

Now we consider the middle range. Here we find the librarian (52.0), real estate agent (52.0), secondary school teacher (51.9), educational counseling director (51.0), social pedagogue (50.5), bookseller (50.5), precision mechanic (49.1), cook (48.6), chemical laboratory technician (47.1), and forester (46.2). The comparison between librarian and forester is significant ($p = .002$), but the effect size is weak at 0.27. The comparison between librarian and social pedagogue just reaches the significance threshold ($p = .050$), but the effect is very weak ($d = 0.17$), and there is no difference at all between the prestige of the social pedagogue and the forester ($p = .438$; $d = 0.01$). The weak to nonexistent differences in the middle range are not surprising. In this group of ten, the values are closer together, with a range of only 5.8 points. Additionally, there is the central point we have already emphasized: one can only be at the very top or bottom if the ratings are highly consistent, meaning that at the extremes, the dispersion is smaller, and with narrower confidence intervals, even small differences become significant, and the effect size is large.

In the middle range, dispersions can be larger, and thus even somewhat larger differences may be statistically insignificant. We illustrate this in Table 5. This shows the means and standard deviations of the 10 occupations with the lowest and highest agreement.

	Consensus low		Consensus high		
	M	s	s	M	
Politician	55.7	28.5	8.4	91.7	Professor (Medicine)
Pastor	63.0	24.3	10.3	88.2	Physician
Real Estate Agent	52.0	22.9	10.7	83.8	Pilot
Student (Bussin. Admin.)	40.9	22.3	11.0	85.6	Ambassador
Ministerial Councillor	70.8	22.1	11.4	87.9	Judge
Professional Soldier	38.4	21.7	11.7	17.7	Worker Ass. Line
Concert Pianist	70.9	20.9	11.9	19.2	Cashier (Superm.)
Homemaker	24.1	20.4	11.9	21.2	Janitor
Sculptor	53.9	20.4	12.5	83.8	Public Prosecutor
Writer	63.4	20.3	12.7	21.9	Telephone Operator

Table 5. Occupations with the highest and lowest standard deviation.

Apart from homemaker ($M = 24.1$), occupations with large dispersion lie in the middle field (38.4 to 70.9). Occupations with the smallest spread are either at the very top (minimum 83.8) or very bottom (maximum 21.9). There is no overlap between the upper seven (low dispersion), middle ten (high dispersion), and lower three (low dispersion). This pattern is already visible in Figure 5; now we have considered concrete occupations instead of dots.

This is now the place to return to the comparisons of medical professor vs. physician and social pedagogue vs. forester. For the frontrunners, professor of medicine and physician, the 3.5-point difference corresponds to an effect size of 0.42. This is solely due to the fact that these two occupations have the smallest standard deviation, so confidence intervals are smallest here. Between the social pedagogue and

forester, who both fall in the middle range, the difference is larger at 4.3 points, but due to greater dispersion, it is statistically meaningless, and the effect size is vanishingly small. These considerations should have made clear that differences must be interpreted with caution.

Sex Differences and Similarities

We now expand the view in two ways. First, we look at the complete list. Second, we differentiate between the judgments of men and women.

In the analysis of rater agreement, we saw that the correlation between the male and female group standards is .972. This is nearly a perfect match. We also noted, however, that correlations pertain to relative ordering. In the following, we show that in absolute rating, there are many similarities between the sexes, but also some profound differences. The primary aim is to show which differences are meaningful and which are not.

Before turning to occupations, we consider in Figure 7 the response styles of men and women. The blue curve shows the smoothed distribution of scores for men, the green for women.

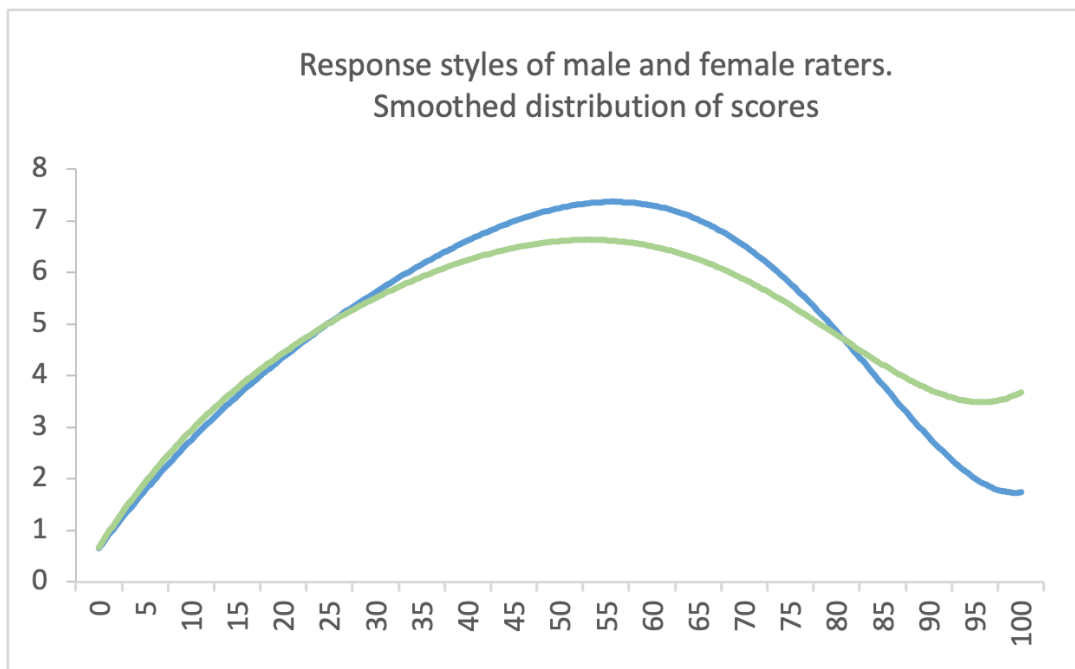


Figure 7. Response styles of male and female raters.

In assigning low ratings, men and women agree, but in assigning very positive ratings, women are much „more generous“ than men. In particular, they assign the maximum 100 much more frequently. The difference between 2.18 and 4.00 percent is significant ($p = .036$), and the effect size is 0.37.

How the different response styles affect occupational ratings is shown in Table 6. Here, occupations are ordered in descending order by prestige in the total group. These values are in the data column to the right of the occupation titles. Beside them are the male and female group standards and the respective difference (M-F: men minus women), both for absolute values and ranks. At the end of the table are statistical descriptors for the individual columns.

Rank	Occupation	Prestige	Males	Females	M-F	M	F	M-F
1	Professor (Medicine)	91.7	90.2	92.8	-2.5	1	1	0
2	Physician	88.2	85.1	90.7	-5.6	2	3	-1
3	Judge	87.9	84.4	90.7	-6.3	3	3	1
4	Ambassador	85.6	83.7	87.1	-3.4	4	5	-1
5	Pilot	83.8	81.4	85.8	-4.4	5	7	-2
6	Public Prosecutor	83.8	78.0	88.4	-10.5	8	4	4
7	Professor (Mathem.)	83.8	80.7	86.2	-5.5	6	6	0
8	Manager (Car Industry)	79.7	79.5	79.9	-0.4	7	10	-3
9	Architect	79.4	74.7	83.2	-8.6	11	8	3
10	Dentist	78.5	72.7	83.1	-10.4	13	9	4
11	Film Director	77.5	75.3	79.3	-3.9	9	11	-2
12	Professor (Art History)	77.0	74.7	78.7	-4.0	10	12	-2
13	Veterinarian	76.1	73.6	78.0	-4.4	12	13	-1
14	Construct. Comp. Owner	74.5	70.6	77.7	-7.1	16	14	2
15	Programmer	73.9	71.3	76.0	-4.7	14	17	-3
16	Graduate Engineer	73.1	69.3	76.1	-6.8	17	16	1
17	Interior Designer	71.2	67.3	74.3	-7.0	18	18	0
18	Concert Pianist	70.9	71.1	70.7	0.3	15	21	-6
19	Ministerial Councillor	70.8	63.4	76.7	-13.3	27	15	12
20	Criminal Inv. Inspector	68.5	64.6	71.6	-7.0	24	19	5
21	Hotel Owner	67.9	64.8	70.4	-5.6	22	22	0
22	Journalist	67.9	63.9	71.1	-7.2	26	20	6
23	Commercial Artist	67.4	65.9	68.6	-2.7	20	24	-4
24	Opera Singer	67.4	67.2	67.5	-0.4	19	25	-6
25	Management Consultant	65.7	64.5	66.7	-2.2	25	26	-1

Rank	Occupation	Prestige	Males	Females	M-F	M	F	M-F
26	Psychologist	65.1	60.4	68.7	-8.3	32	23	9
27	Actor (Theater)	64.9	65.8	64.1	1.7	21	30	-9
28	Savings Bank Manager	64.0	61.9	65.6	-3.7	30	28	2
29	Tax Consultant	63.7	60.4	66.3	-5.8	33	27	6
30	Writer	63.4	63.3	63.5	-0.2	29	32	-4
31	Pastor	63.0	61.4	64.3	-2.9	31	29	2
32	Broadcaster	62.7	64.6	61.1	3.5	23	33	-10
33	Goldsmith	59.7	63.3	56.9	6.3	29	38	-10
34	Sports Reporter	59.2	58.4	59.9	-1.5	35	34	1
35	Translator	58.7	58.3	59.0	-0.7	36	35	1
36	Firefighter	56.0	58.6	54.0	4.6	34	44	-10
37	Primary School Teacher	56.0	54.0	57.6	-3.6	40	36	4
38	Camera Operator (TV)	55.9	55.3	56.3	-1.1	37	40	-3
39	Bank Employee	55.8	53.6	57.6	-3.9	42	37	5
40	Politician	55.7	45.6	63.7	-18.1	59	31	28
41	Photographer	55.0	53.9	55.8	-1.9	41	42	-1
42	Draftsman	53.9	51.6	55.8	-4.2	46	41	5
43	Food Chemist	53.9	54.7	53.3	1.4	38	47	-9
44	Sculptor	53.9	54.6	53.3	1.3	39	46	-7
45	Librarian	52.0	49.0	54.4	-5.4	48	43	5
46	Real Estate Agent	52.0	46.0	56.8	-10.8	58	39	19
47	Second. School Teacher	51.9	49.7	53.5	-3.8	47	45	2
48	Educ. Counseling Direct.	51.0	48.6	52.8	-4.2	50	49	1
49	Social Pedagogue	50.5	47.1	53.2	-6.1	55	48	7
50	Bookseller	50.5	51.8	49.4	2.3	45	50	-5
51	Precision Mechanic	49.1	52.0	46.8	5.2	44	51	-7

Rank	Occupation	Prestige	Males	Females	M-F	M	F	M-F
52	Cook	48.6	52.4	45.6	6.8	43	54	-11
53	Chemical Laborat. Techn.	47.1	47.6	46.7	0.8	52	52	0
54	Forester	46.2	48.9	44.0	4.8	49	59	-10
55	Nurse	45.3	47.4	43.6	3.8	53	61	-8
56	Opinion Researcher	44.8	46.0	43.8	2.2	57	60	-3
57	Tour Guide	44.6	43.9	45.1	-1.2	62	56	6
58	Food Wholesaler	44.3	43.5	44.9	-1.4	63	57	6
59	Jewelry Store Salesman	44.2	48.2	41.1	7.1	51	72	-21
60	Retail Salesman	44.1	45.0	43.5	1.5	60	62	-2
61	Druggist	44.0	46.9	41.7	5.1	56	70	-14
62	Art Teacher	43.5	41.1	45.4	-4.3	69	55	14
63	Traffic Policeman	43.5	39.5	46.7	-7.1	73	53	20
64	Administrative Employee	43.4	42.0	44.4	-2.4	68	58	10
65	Innkeeper	43.1	43.3	43.0	0.3	64	64	0
66	Baker	43.0	47.3	39.7	7.7	54	76	-22
67	Auto Mechanic	42.4	43.2	41.8	1.4	65	68	-3
68	Educator (Kindergarten)	42.0	43.1	41.3	1.8	66	71	-5
69	Payroll Accountant	41.9	40.6	42.9	-2.3	70	65	5
70	Train Driver	41.5	44.1	39.5	4.6	61	77	-16
71	Student (Business Adm.)	40.9	39.0	42.3	-3.3	76	66	10
72	Accountant	40.7	39.3	41.8	-2.5	74	67	7
73	Office Worker	40.7	39.3	41.8	-2.5	75	69	6
74	Dispatch Manager	40.2	39.9	40.4	-0.6	72	74	-2
75	Finance Officer	40.0	35.6	43.4	-7.8	79	63	16
76	Dance Teacher	39.6	40.4	39.0	1.4	71	78	-7
77	Decorator	38.8	37.4	39.8	-2.4	77	75	2

Rank	Occupation	Prestige	Males	Females	M-F	M	F	M-F
78	Professional Soldier	38.4	35.5	40.8	-5.2	80	73	7
79	Master Painter	38.4	42.7	35.1	7.6	67	80	-13
80	Commercial Agent	36.1	35.2	36.7	-1.5	81	79	2
81	Farmer	34.7	34.8	34.6	0.2	82	81	1
82	Butcher	32.9	36.6	30.0	6.6	78	85	-7
83	Hairdresser	31.8	33.3	30.6	2.7	83	83	0
84	Waiter	30.2	31.8	28.9	2.8	84	86	-2
85	Postal Clerk	30.0	28.8	31.0	-2.3	90	82	8
86	Bus Driver	29.1	30.2	28.2	2.1	85	87	-2
87	Insurance Agent	28.8	27.0	30.2	-3.2	92	84	8
88	Warehouse Manager	28.1	29.1	27.3	1.8	89	88	1
89	Truck Driver	27.8	29.7	26.2	3.5	88	89	-1
90	Deptmt. Store Salesman	27.3	29.9	25.3	4.5	86	92	-6
91	Lifeguard	27.2	29.8	25.2	4.6	87	93	-6
92	Cab Driver	26.0	26.4	25.8	0.6	93	91	2
93	Gas Station Attendant	24.6	28.7	21.3	7.4	91	94	-3
94	Homemaker	24.1	21.8	26.0	-4.2	96	90	6
95	Telephone Operator	21.9	22.7	21.3	1.4	95	95	0
96	Janitor	21.2	23.0	19.8	3.1	94	96	-2
97	Cashier (Supermarket)	19.2	20.4	18.3	2.2	98	97	1
98	Worker (Assembly Line)	17.7	20.6	15.4	5.2	97	98	-1
	Minimum	17.7	20.4	15.4	-18.1			
	Maximum	91.7	90.2	92.8	7.7			
	Median	50.5	48.7	51.1	-1.7			
	Mean	52.1	51.3	52.7	-1.5			

Rank	Occupation	Prestige	Males	Females	M-F	M	F	M-F
	Standard deviation	18.2	17.0	19.5	4.9			

Table 6. Comparison men/women. Absolute values and ranks. (*italics=red; bold=green*)

Overall, women's ratings are 1.5 points higher than men's (52.7 vs. 51.3). That's not a lot, but the difference is significant ($p = .005$) and the effect size is 0.29. However, the differences are not evenly distributed across the prestige spectrum.

In addition to the group standards, the M-F column shows the difference between men and women. Positive values indicate higher ratings by men, while negative values indicate higher ratings by women. Differences of more than 5 points are highlighted in green for higher ratings by women and red for higher ratings by men. Two points immediately stand out. First, the largest differences have a negative sign, meaning that women gave much higher ratings. Second, there are many green marks in the upper half, but only one red. A count shows 42 negative signs and only 7 positive signs in the upper half. This means that women give much more favorable ratings than men in this section. On average, the ratings are 4.1 points higher. In the lower half of the table, the differences are considerably smaller. Here, men give better ratings. On average, however, the difference is only 1.2 points. The sex difference becomes even more pronounced when looking at the top 10 and bottom 10 occupations. The top group is rated 5.8 points better by women than by men (86.8 vs. 81.0), while the bottom group is rated 2.8 points better by men (25.3 vs. 22.5).

The sex difference is reflected in the dispersion. Among women, we see a stronger differentiation between occupations. The minimum is 15.4 and the maximum is 92.8. Among men, the values range from 20.4 to 90.2. This means that the range among women is 7.6 points greater (77.4 vs. 69.8). The standard deviation is 19.5 for women and 17.0 for men.

The three rightmost columns show the rank positions. Since rank positions are inversely related to the point scores, a negative sign indicates a higher, thus worse, rank position in the eyes of men. The most striking result is found at the top. Among the first 17 occupations, there are 9 differences of more than 5 points (all with higher ratings by women), but the differences in rank order are minimal. The largest rank difference is 4 positions, occurring twice, namely for the public prosecutor and the dentist, who are rated

10.5 and 10.4 points higher, respectively. This example shows that even substantial absolute differences can be accompanied by only minor shifts in rank order. A counterexample is the food chemist and the actor (theater); here, the difference is only 1.4 and 1.7 points, respectively, but each has a rank difference of 9 positions. Furthermore, there are several seemingly curious constellations. For the salesperson in the jewelry store, the traffic cop, and the owner of a construction company, the difference is 7.1 points in each case, and for the interior designer, it is 7.0. For the first two, the rank difference is 21 and 20, but for the second two, it is 2 and 0. For the train driver, a difference of 4.6 points corresponds to a rank difference of 16; for the hotel owner, the difference is one point larger (5.6), but the rank position is the same for men and women (22). The architect (8.6) has a larger point difference than the baker (7.7), but the rank difference for the architect is only 3, while for the baker, it is 22. These examples vividly underscore that point differences and rank differences can diverge in surprising ways in some cases. This should once again make it clear that differences, whether absolute or relative, must be interpreted with caution. The politician is also worth mentioning. Women rate the politician much higher than men (63.5 vs. 45.6). Thus, the politician exhibits not only the largest absolute difference but also the largest relative difference (rank 31 vs. 59). The second-largest point difference is found for the ministerial councillor (women: 76.7; men: 63.4), a political official, but the rank difference is „only” 27 versus 15.

As a key finding, it should be noted that the difference between male and female raters is related to occupational prestige, with women rating high-prestige occupations substantially better than men, while men rate lower-prestige occupations slightly better than women. We now examine this highly interesting relationship from a different perspective. For each occupation, we calculated the significance and effect size d for the sex difference using a t -test. Table 7 shows occupations where the difference is significant at the 1% level. In addition to the total group prestige values, the table shows the difference between male and female raters ($M-F$), the p -value, and the effect size. A positive difference ($M-F$) and a positive effect size mean that men rated the profession higher, while a negative value indicates a higher rating by women.

Occupation	Prestige	M-F	p	Effect size
Public Prosecutor	83.8	-10.5	< .001	-0.91
Architect	79.4	-8.6	< .001	-0.70
Dentist	78.5	-10.4	< .001	-0.69
Politician	55.7	-18.1	< .001	-0.67
Ministerial Councillor	70.8	-13.3	< .001	-0.62
Judge	87.9	-6.3	.001	-0.57
Physician	88.2	-5.6	.002	-0.56
Graduate Engineer	73.1	-6.8	.003	-0.52
Interior Designer	71.2	-7.0	.004	-0.51
Real Estate Agent	52.0	-10.8	.006	-0.48
Criminal Investig. Inspector	68.5	-7.0	.007	-0.47
Construction Company Owner	74.5	-7.1	.008	-0.46
Worker (Assembly Line)	17.7	5.2	.010	0.45
Baker	43.0	7.7	.009	0.46
Master Painter	38.4	7.6	.007	0.47
Gas Station Attendant	24.6	7.4	.002	0.54

Table 7. Effect size of significant sex differences ($p \leq .01$).

From this perspective, our central finding is immediately apparent: The differences between men and women primarily concern the upper range, where women assign much more positive ratings than men. The sex differences listed here are not only statistically significant but also exhibit medium to very high effect sizes. Furthermore, the table confirms another point we have highlighted: For some occupations—worker (assembly line), physician, judge, graduate engineer—the point difference is not very large, but due to the low dispersion, it is statistically significant, and the effect size is substantial.

Figure 8 illustrates the relationship between occupational prestige and the effect size of the sex difference for all occupations.

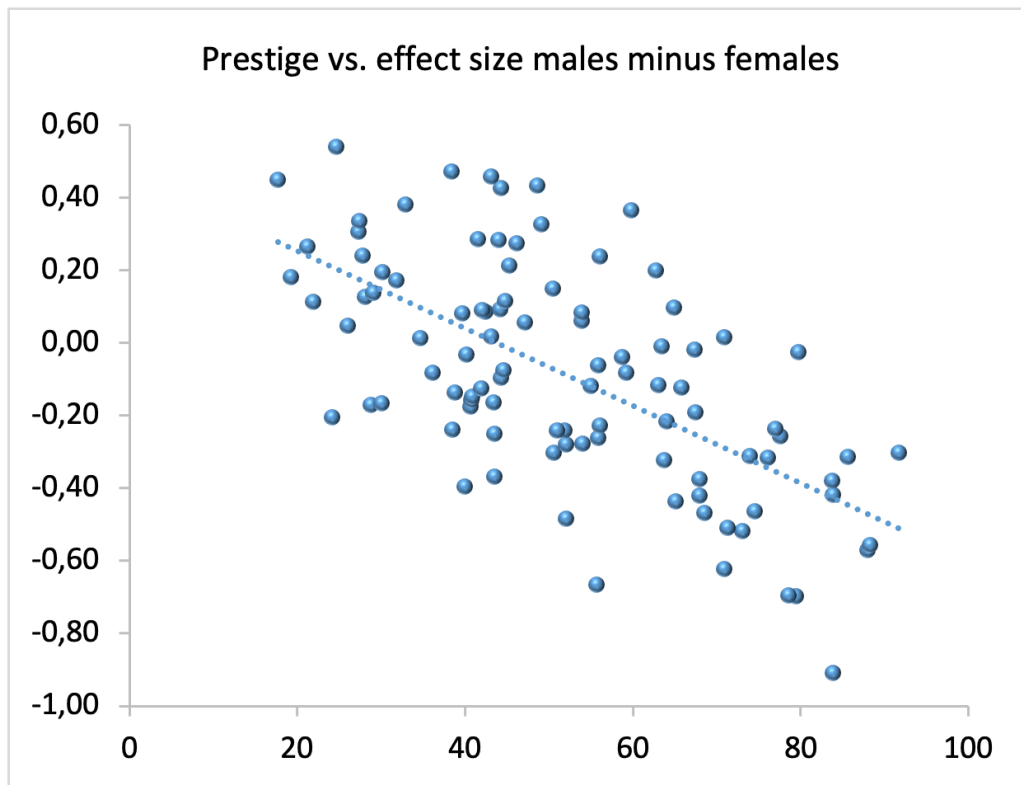


Figure 8. Prestige vs. effect size males minus females.

The picture is unequivocal: The relationship between occupational prestige and the effect size of the sex difference is linear, and since we consider the men minus women the difference, it is negative. The correlation is $-.635$; this corresponds to a strong effect.

To conclude this section, a glance at the beginning is warranted. There, we first noted that the correlation between the male and female group standards is $.972$. Thus, in relative terms, there is nearly perfect agreement between the sexes, but within these commonalities, we have shown systematic differences and some astonishing relationships between absolute differences, ranks, statistical significance, and effect size. In the occupational prestige literature, this is rarely discussed.

West Germany, East Germany

Due to its historical development, Germany presents a peculiarity. After World War II, Germany was divided into two blocs that stood irreconcilably opposed to each other during the Cold War. Since 1990, the

Federal Republic of Germany has been a united nation, but the traces of decades of division are still present today. This is especially true for the year 1999, in which our data collection took place. We therefore also highlight the differences between East and West.

Here, the problem arises that only 13 participants are from East Germany. Since the group standard exhibits high reliability (.927) and the gender ratio is balanced, a comparison is justified despite the very small sample size.¹⁶

The correlation between the group standards is .961. Thus, East and West Germany show an extraordinarily high level of agreement in relative ratings. This also applies to the average rating across all occupations, which is 51.9 in the West and 52.2 in the East. In contrast, there are significant differences in dispersion. In West Germany, the range spans from 17.1 to 92.1, while in East Germany, it spans from 21.7 to 90.7. The standard deviation is 18.6 versus 15.8.

For the vast majority of occupations, there is great agreement between East and West; the main differences are summarized in Table 8. This table shows the occupations with a difference of at least 7 points, with a complete list provided in Table A3 in the appendix. The red-marked positive values indicate higher ratings by West German participants, while the green-marked negative values indicate higher ratings by East German participants. The occupations are listed in descending order based on their prestige in West Germany.

Occupation	West	East	W-E
Manager (Car Industry)	81.9	71.2	10.7
Architect	80.2	69.9	10.3
Professor (Art History)	79.2	69.2	10.0
Dentist	79.0	71.2	7.8
Film Director	78.8	70.1	8.7
Ministerial Councillor	72.3	63.8	8.5
Concert Pianist	71.4	60.5	10.9
Hotel Owner	69.0	61.2	7.8
Commercial Artist	67.5	59.5	8.0
Pastor	65.1	57.8	7.3
Goldsmith	61.3	52.1	9.2
Firefighter	54.8	62.2	-7.4
Draftsman	53.1	60.5	-7.4
Forester	48.2	38.9	9.3
Chemical Laboratory Technician	45.1	54.2	-9.1
Administrative Employee	42.7	50.6	-7.9
Retail Salesman	42.7	49.9	-7.2
Payroll Accountant	40.8	53.1	-12.3
Accountant	39.9	49.2	-9.3
Educator (Kindergarten)	39.4	47.0	-7.6
Dance Teacher	39.0	48.5	-9.5
Finance Officer	38.7	49.5	-10.8
Department Store Salesman	26.2	37.6	-11.4
Truck Driver	26.1	35.2	-9.1
Cashier (Supermarket)	18.1	28.1	-10.0

Table 8. Marked differences between West and East Germany. (*italics=red; bold=green*)

The picture could hardly be clearer. The occupations that are rated markedly higher in the West than in the East enjoy high prestige; those that perform better in the East have much lower prestige. Using the entire group of 136 participants as a benchmark, the average prestige in the upper group is 73.0, and in the lower group, it is 38.5. Even within these two groups, there is a clear tendency for the discrepancy to increase toward the upper or lower end of the prestige scale. Across all occupations, the correlation between prestige and the West-East difference is .591.

The crucial point is as follows. The table includes the occupations where the ratings between East and West diverge most in absolute terms, but both agree on the relative order, with a correlation of .961. This means that the East German participants perceive practically the same hierarchical sequence, but they have significantly leveled the differences. This is undoubtedly a consequence of over forty years of communist regime rule, which, even in 1999—a decade after its end—left a deep imprint.

Male Occupations, Female Occupations

Thus far, we have focused on the 1999 main study, in which occupations were rated only with reference to men. Now we also consider the 1990 and 1997 surveys, opening a new horizon, as there occupational titles were considered both in male and female form. Since there is no simple term for this distinction, for the sake of simplicity, we will refer to male occupations and female occupations, with the explicit note that this does not refer to the relative proportion of the sexes.

In 1990, the list included 95 occupations. Except for professional soldier and secretary, the male and female versions referred to the same occupation. The 95 male occupations were all included in the 1999 main study. The means of male and female occupations can be found in the appendix in Table A2.

The correlation between men's and women's occupations is .984. The minimum is 5.94 in the male version and 8.32 in the female. The maximum is 91.84 and 90.34. Means (48.23 vs. 48.12) and standard deviations (21.39 vs. 21.49) are nearly identical. Overall, there is overwhelming agreement, but some details should be mentioned here.

The 10 occupations with the largest score difference in favor of men are: firefighter (10.5), precision mechanic (9.0), pilot (9.0), forester (8.2), pastor (7.7), office worker (6.9), butcher (6.8), manager (car

industry) (6.0), auto mechanic (5.5), and film director (4.4). The 10 occupations where women are attributed higher prestige are: homemaker (12.3), nurse (7.6), actress (theater) (6.3), dance teacher (6.1), social pedagogue (6.0), decorator (5.4), primary school teacher (5.2), art teacher (5.0), opera singer (4.5), and opinion researcher (4.1). Most differences are negligible. However, another point is noteworthy: in occupations where men have higher prestige than women, the proportion of women is very small, except for office workers; conversely, in occupations where women perform better, the proportion of women is significantly larger, and in some cases, these are typical female-dominated domains. Overall, the differences are remarkably small, but there is a tendency for men and women to be relatively higher regarded in some occupations where they are more strongly represented.

The comparison with the main study of 1999 shows the following: For men's occupations, the test-retest after 9 years with a completely different rater sample yields a correlation of .974. Even the comparison of men's occupations from 1999 with women's occupations from 1990 yields a correlation of .953.

It should be reiterated that the raw data from 1990 are no longer available. However, the extraordinarily high rater agreement in the 1999 and 1997 ratings and the extremely high correlations between the group standards of 1990 and 1999 make it clear that the means still available to us must have extremely high reliability.

In the two sub-experiments on multiple male choice criteria from 1997, the prestige of 40 men's and 40 women's occupations was assessed. Of these, 34 referred to the same occupation. The correlation between the two variants is, as shown in Table 3, .970.

For the 34 common occupations, the mean for the male variant is 57.5, and for the female variant, 56.0. The top 10 occupations where men enjoy higher prestige than women are: commercial artist (10.6), business administrator (10.5), dispatch manager (8.4), bank employee (8.3), programmer (8.0), mathematician (7.1), finance officer (5.9), lawyer (5.8), druggist (4.7), and retail salesman (4.2). A notable reversal occurs only for the actress and the broadcaster, whose prestige is rated 8.6 and 6.7 points higher, respectively, than that of their male counterparts. These are followed by chemical laboratory technician (3.7), lifeguard (3.5), psychologist (3.3), goldsmith (3.0), decorator (2.7), primary school teacher (2.1), telephone operator (2.1), and waitress (1.8). This confirms the trend from the 1990 study. It should be noted that in 1997, men's occupations were assessed only by women, and women's occupations only by men. In 1990, both sexes evaluated the occupations.

Here, too, we consider the comparison with the main study of 1999. With women's occupations, there are 32 overlaps, with a correlation of .853. For men's occupations, there are 31 overlaps, with a correlation of

.887.

Overall, our comparisons make it clear that prestige depends much more on the occupation than on the sex of the occupational holder. However, there is also a tendency for sexes to be rated relatively more favorably in occupations where their representation is higher.

Changes from 1990 to 1999

Now we compare the prestige ratings of male occupations from 1990 and 1999. As already stated, the correlation is .974. Nonetheless, there are some noteworthy changes here as well. Table 9 shows in the upper half the ten occupations with the largest absolute gains and in the lower half those with the largest losses.

	Points			Rank		
Occupation	1999	1990	1999-1990	1999	1990	1990-1999
Innkeeper	43.1	23.2	20.0	64	82	18
Firefighter	56.0	37.2	18.8	35	63	28
Waiter	30.2	16.7	13.5	81	87	6
Cook	48.6	35.5	13.1	51	66	15
Gas Station Attendant	24.6	11.6	13.0	90	93	3
Worker (Assembly Line)	17.7	5.9	11.8	95	95	0
Homemaker	24.1	12.7	11.4	91	92	1
Accountant	40.7	29.6	11.2	69	76	7
Department Store Salesman	27.3	16.4	11.0	87	89	2
Baker	43.0	32.3	10.7	65	75	-10
Professor (Mathematics)	83.8	86.8	-3.0	6	3	-3
Graduate Engineer	73.1	76.2	-3.1	15	13	-2
Forester	46.2	49.6	-3.5	53	45	-8
Construction Comp. Owner	74.5	78.9	-4.3	13	11	-2
Pastor	63.0	67.7	-4.6	30	20	-10
Ministerial Councillor	70.8	76.3	-5.5	18	12	-6
Veterinarian	76.1	81.7	-5.6	12	8	-4
Dentist	78.5	86.0	-7.5	9	5	-4
Finance Officer	40.0	47.9	-7.9	72	46	-26
Politician	55.7	63.7	-8.0	39	25	-14

Table 9. Changes in occupational prestige between 1990 and 1999.

The largest gains are seen for the innkeeper (20.0) and the firefighter (18.8). For the innkeeper, this represents an improvement of 18 ranks, and for the firefighter, an even greater improvement of 28 ranks. These are substantial improvements in both respects. In contrast, the gas station attendant, department store salesman, homemaker, and worker on the assembly line show gains of 13.0, 11.0, 11.4, and 11.8 points, respectively, but the gas station attendant improves by only three ranks, the department store salesman by two, the homemaker by one, and the worker on the assembly line remains at the bottom. At the lower end, occupations can indeed enjoy considerable absolute gains, but they still rank at the bottom.

Comparing the upper and lower halves of the table makes it clear that point losses are much smaller than gains. The largest losses are suffered by the politician (-8.0) and the finance officer (-7.9). The politician drops 14 ranks, and the finance officer drops 26 ranks. Although the pastor loses fewer points than the dentist (-4.6 vs. -7.5), the pastor loses 10 ranks, while the dentist loses only 4. From the middle range, two further cases are worth mentioning. The train driver gains 1.9 points but loses 10 ranks (from 58 to 68), and the master painter gains 0.6 points but loses 15 ranks (from 61 to 76). In both cases, slight absolute gains are accompanied by substantial relative losses.

Table 10 shows correlations between points, ranks, and differences in 1990 and 1999.

	Points			Rank	
	P99	P90	P99-90	R99	R90
P1990	.974				
P1999-1990	-.489	-.675			
R1999	-.991	-.963	.478		
R1990	-.964	-.993	.680	.964	
R1999-1990	-.099	.114	-.760	.133	-.133

Table 10. Correlations between points, ranks, and differences 1990, 1999.

Within the two surveys, the correlation between scores and ranks is, as expected, nearly perfect, with

values of .993 in 1990 and .991 in 1999. The comparison between 1990 and 1999 also shows an extraordinarily high level of agreement for both scores and ranks, at .974 and .964, respectively. This indicates that the prestige hierarchy remains almost identical over this time period (and across different rater samples and settings). This confirms what has long been known from the literature; and it does not matter whether absolute values or ranks are considered. A new insight emerges from a different perspective. The correlation between the score difference (1999 minus 1990) and the rank difference (1999 minus 1990) is $-.760$. The negative sign necessarily arises because scores and ranks are inversely related. The magnitude is noteworthy. The relationship is still very strong but far from perfect. This is not surprising, as differences are considerably less reliable than individual measurement series. The interesting point lies elsewhere. The rank changes from 1999–1990 show neither a correlation with the scores from 1990 (.114) nor with the scores from 1999 ($-.099$), and they are also not related to the ranks (1990: $-.133$; 1999: .133). A completely different picture emerges for the score difference 1999 minus 1990. The correlation with the scores from 1999 is $-.489$, and with the scores from 1990, it is even $-.675$. This reflects the finding that we highlighted above: in 1999, occupations in the lower range were rated much better than in 1990. For example, the minimum rose from 5.94 to 17.7, but nevertheless, the assembly line worker still stood in last place in 1999, and some other occupations, although they gained considerably in prestige in absolute terms, still ranked far down in 1999.

Thus, from a temporal perspective, we observe that absolute and relative differences are naturally closely related. However, in some cases, they diverge in surprising ways and can even be contradictory.

Prestige of One's Own Occupation and External Judgment

At the end of the survey, participants were asked to indicate the prestige of their own occupation. The ratings range from 0 to 100, with an interquartile range of 40 to 70. Of the participants who rated their own occupation below 40, 15 were still in training, 6 were office workers, and one each identified as a homemaker, gardener, farmer, geriatric nurse, toolmaker, and casting director, with one participant providing no information. On average, the prestige of one's own occupation is 54.2. This is 2.16 points higher than the average prestige of the 98 occupations, but the difference is not significant ($p = .292$). For completeness, it should be noted that the self-assessed prestige of one's own occupation correlates with the average rating of all occupations ($r = .209$; $p = .015$). However, it seems trivial that those who generally give higher ratings also rate themselves higher, especially since the correlation is only weak.

Intra-German Changes Over Time

So far, we have only considered our own studies. Now, we broaden our perspective and compare our results with other studies. Initially, we remain focused on Germany, where we can span a timeline from 1952 to 2018.

Our attention is directed to the two most significant German occupational prestige scales, namely the Magnitude Prestige Scale (MPS) by Wegener^[35] and the Berufliche Ansehensskala (BAS) by Ebner and Rohrbach-Schmidt^[38].

The MPS is based on surveys conducted in 1979 and 1980 and is distinguished by its use of magnitude scaling, a method primarily rooted in psychophysics. This approach results in a right-skewed distribution, meaning that low prestige values are compressed, while high prestige values are widely dispersed. This is illustrated in Figure 9.

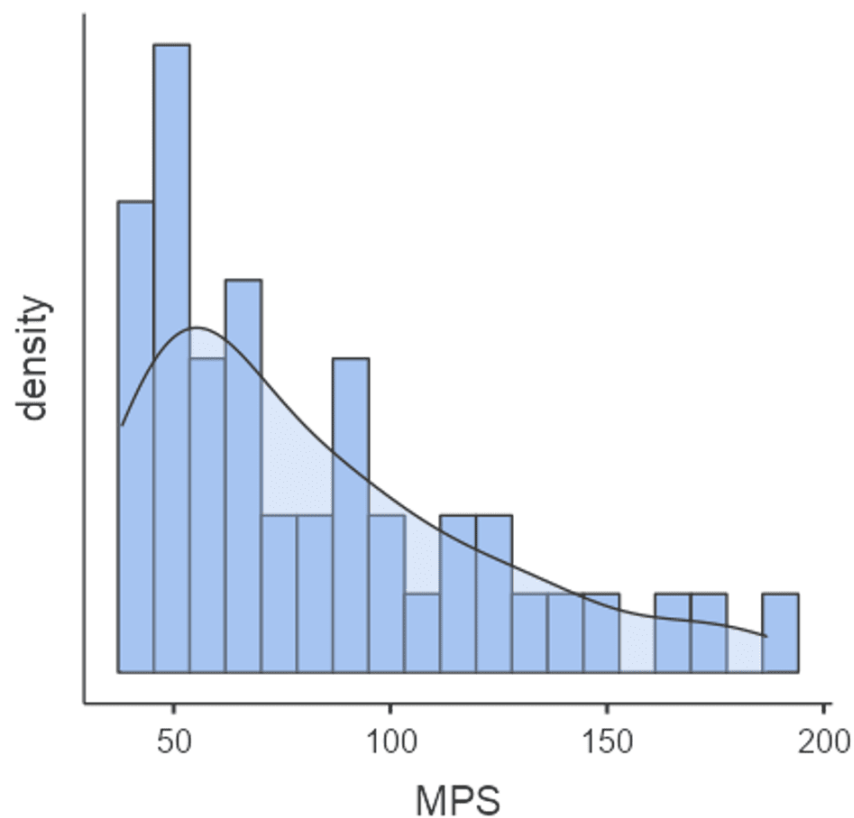


Figure 9. Distribution of MPS scores.

We logarithmized the scale, resulting in correlations that are .02 to .06 higher in several of the following comparisons. We report the correlations with the logarithmized version in the following, without mentioning this each time.¹⁷

The surveys for the BAS were conducted in 2017/2018. Fortunately, the ratings were made on a scale from 0 to 10, making them comparable to ours through a factor of 10.

Our surveys from 1990 and 1999 fit well into this framework, with successive measurement points separated by 10, 9, and 19 years, and the main study from 1999 situated in the middle. Through a critical matching based on occupational titles, 77 of our occupations are found in the BAS and 44 in the MPS. Of these, 41 are represented in all lists.

In addition to the MPS and BAS, we also consider the earlier studies by Bolte^[33] and Kleining and Moore^[34], whose data collection took place in the early and late 1950s, respectively. Of the 38 occupations rated by Bolte's urban sample, 17 are found in our list,¹⁸ and the intersection with the 70 occupations from Kleining and Moore amounts to 28. In Table 11, we show only the correlations with our 1990 and 1999 scales and between the MPS and BAS, as the other intersections are too small. It should be noted that the values are not directly comparable due to the different samples of occupations.

	1999	1990	MPS 1980
1990	.974		
MPS 1980	.844	.849	
BAS 2018	.758	.724	.567
Bolte 1952	.858	.881	
K&M 1958	.891	.896	

Table 11. Correlations between German occupational prestige ratings 1952 to 2018.

The closest relationship is shown between our 1990 and 1999 scales and the survey by Kleining and Moore from 1958 (.896 and .891) and Bolte's survey from 1952 (.881 and .858). The MPS from 1980 follows closely behind (.844 and .849). In contrast, the correlation with the BAS drops significantly (.724 and .758). The

correlation between the MPS and BAS is even lower at .567. This is not due to our small intersection, as Ebner and Rohrbach-Schmidt^[38] report a correlation of only .62 for 135 common occupations between their BAS and the MPS. Thus, it is beyond doubt that our scales exhibit a much closer alignment with the MPS than with the BAS, and that the BAS is an outlier.

To gain insights into changes in individual occupations, we compared the successive surveys. Due to the small intersection, data from the 1950s are not considered. Since the scales have different means and standard deviations, we z-transformed them. For each occupation, the z-score of the earlier survey was subtracted from the z-score of the subsequent survey. Table 12 lists the occupations with an observed difference of at least 0.75 standard deviation units.¹⁹ A positive difference indicates that the occupation improved its relative position in the respective period, while a negative difference indicates a decline. For orientation, the first data column provides the score values from our main study in 1999, which is temporally situated in the middle of the observation period.

		z-score difference		
	1999	1990 - MPS	1999 - 1990	BAS - 1999
Firefighter	56.0	0.10	0.73	1.61
Traffic Policeman	43.5		0.13	1.53
Baker	43.0	0.26	0.24	1.39
Auto Mechanic	42.4	0.25	0.11	1.14
Butcher	32.9	-0.10	-0.04	1.03
Warehouse Manager	28.1	0.10	-0.11	0.98
Bus Driver	29.1		0.04	0.95
Food Wholesaler	44.3	-0.84	-0.18	0.85
Educator (Kindergarten)	42.0	-1.95		0.82
Criminal Investig. Inspector	68.5	0.96	0.12	0.79
Train Driver	41.5	0.65	-0.18	0.79
Chemical Laboratory Techn.	47.1		0.16	0.79
Tour Guide	44.6	1.33	-0.32	
Goldsmith	59.7	0.80	-0.07	0.06
Commercial Agent	36.1	-1.18	0.01	-0.47
Insurance Agent	28.8	-1.60	-0.12	-0.27
Concert Pianist	70.9		-0.06	-0.75
Politician	55.7		-0.52	-0.79
Actor (Theater)	64.9	-0.17	0.26	-0.80
Telephone Operator	21.9	-0.65	-0.05	-0.82
Journalist	67.9		-0.10	-0.91
Druggist	44.0		-0.03	-0.95
Cook	48.6	0.67	0.40	-1.04
Commercial Artist	67.4		0.36	-1.12

		z-score difference		
	1999	1990 - MPS	1999 - 1990	BAS - 1999
Film Director	77.5	-0.09	0.22	-1.46
Management Consultant	65.7		0.27	-1.49
Pastor	63.0	0.45	-0.30	-2.00

Table 12. Changes between consecutive surveys. (*italics=red; bold=green*)

One result immediately stands out: In our own surveys from 1990 and 1999, there was no shift of three-quarters of a standard deviation. However, the firefighter narrowly missed this threshold (0.73). With regard to later considerations, it should be noted that the politician lost prestige, though not dramatically (-0.52).

In comparing the Magnitude Prestige Scale from 1980 with our 1990 survey, the tour guide (1.33), the criminal investigation inspector (0.96), and the goldsmith (0.80) significantly improved their relative position. On the other hand, the educator (kindergarten) (-1.95), the insurance agent (-1.60), the commercial agent (-1.18), and the food wholesaler (-0.84) suffered losses in prestige.

Particularly interesting is the comparison of our 1999 scale with the BAS from 2018. Here, a clear division emerges. In 2018, the firefighter (1.61), the traffic policeman (1.53), the baker (1.39), the auto mechanic (1.14), the butcher (1.03), the warehouse manager (0.98), the bus driver (0.95), the food wholesaler (0.85), the educator (kindergarten) (0.82), the criminal investigation inspector (0.79), the train driver (0.79), and the chemical laboratory technician (0.79) achieved a significantly better relative position. With the exception of the criminal investigation inspector and the firefighter, these occupations were in the lower half in 1999. On the losing side are the pastor (-2.00), the management consultant (-1.49), the film director (-1.46), the cook (-1.04), the druggist (-0.95), the journalist (-0.91), the telephone operator (-0.82), the actor (theater) (-0.80), the politician (-0.79), and the concert pianist (-0.75). With the exception of the cook, the druggist, and the telephone operator, these occupations were in the upper half in 1999, but in 2018, some are found in the lower half. Thus, the BAS shows a strong leveling effect. For the 77 common occupations, there is a significant correlation of -0.345 ($p = .002$) between the z-scores from 1999 and the change in z-scores.

German Prestige Scales in International Comparison

We now compare the German scales with several international scales. There, we have, on the one hand, the Standard International Occupational Prestige Scale (SIOPS) by Treiman^[1]. This is the classic scale par excellence and the only truly international one in the narrow sense. Our other comparisons relate to China, Great Britain, Poland, Sweden, Spain, South Korea, Hungary, and the USA. In addition to the very old studies by Counts^[7] and Smith^[8], we consider the studies of the National Opinion Research Center (NORC) from 1989 and 2012, the Chinese survey from 1994^[21], the Polish CBOS 2008^[20], the Swedish surveys from 2002 and 2018^[25], the Spanish one from 2013^[24], the Hungarian one from 2016^[28],²⁰ and the online surveys from 2016 in the USA^[26], 2022 in Great Britain (Newlands and Lutz, 2024), and in South Korea 2022^[29]. The ratings in Great Britain pertained, on the one hand, to prestige and, on the other hand, to social value. Furthermore, we consider the study by de Vries et al.^[32], which stands out as an outlier, as it involved no human raters; instead, the assessments were made using artificial intelligence. In addition to prestige, status, boast-worthiness, and training/pay were also considered.

In international comparisons, the matching of occupations constitutes a central problem. Here, we proceeded conservatively and only considered occupations for which we are reasonably certain. This comes at the expense of sample size, but we were able to include occupations from the most important main groups and cover the full breadth of occupational prestige. The number of common occupations is specified in Table 13. Since the temporal comparison is the focus, we indicate the respective survey time point (with the exception of SIOPS 1970, which is based on studies from the 1950s and 1960s).

		MPS 1980	1990/1999	BAS 2018
USA	Counts 1925	13	18	17
USA	Smith 1939-1941	18	31/32	30
International	SIOPS 1970	28	46/47	45
USA	NORC 1989	27	43	41
China	Chiu 1994	23	36/37	35
USA	NORC 2010	34	63	51
Poland	CBOS 2008	10	18/19	19
Spain	CIS 2013	32	52/54	53
Sweden	UE & N 2002	23	32/34	30
Sweden	UE & N 2018	23	32/34	30
USA	Hughes et al. 2016	30	49	46
Hungary	Sik et al. 2016	26	42/44	40
Great Britain	Newlands and Lutz 2022	34	54	47
South Korea	Kye und Soel 2022	11	19	18
AI (USA)	De Vries et al. 2024	20	36	35

Table 13. Number of common occupations.

Most overlaps inevitably occur with our own studies, as we limit our analysis to our own list. If the complete MPS and BAS were considered, they would have a much larger intersection. For the BAS, the number of common occupations is almost as large as in our studies, allowing for a direct comparison. In contrast, significantly fewer overlaps are found with the MPS. The correlations between the German and international scales are presented in Table 14. The studies are sorted row- and column-wise according to their temporal sequence. The variables marked with an asterisk in the last three rows are only more or less closely related correlates of occupational prestige. The MPS/SIOPS comparison is excluded because the MPS is derived from the SIOPS.

		MPS	Own Scales		BAS
		1980	1990	1999	2018
USA	Counts 1925	.87	.92	.91	.66
USA	Smith 1943	.91	.86	.84	.66
International	SIOPS 1970		.88	.84	.60
USA	NORC 1989	.94	.91	.90	.70
China	Chiu 1994	.85	.82	.79	.54
Schweden	UE & N 2002	.93	.93	.95	.70
USA	NORC 2010	.86	.84	.82	.69
Poland	CBOS 2008	.27	.11	.25	.57
Spain	CIS 2013	.51	.58	.65	.79
USA	Hughes et al. 2016	.85	.92	.92	.80
Hungary	Sik et al. 2016	.82	.88	.89	.73
Schweden	UE & N 2018	.91	.90	.94	.79
Great Brit.	N & L 2022 Prestige	.84	.86	.89	.77
South Korea	Kye und Soel 2022	.91	.92	.90	.72
KI (USA)	de Vries et al. 2024 Status	.90	.91	.93	.70
KI (USA)	de Vries et al. 2024 Prest.	.90	.92	.93	.68
Great Brit.	N & L 2022 Soc. Value *	.60	.54	.57	.64
AI (USA)	de Vries et al. 2024 B-w *	.63	.71	.80	.58
AI (USA)	de Vries et al. 2024 T/P *	.59	.61	.67	.68

Table 14. Correlation between German and international scales.

Three scales stand out as outliers, particularly the Polish CBOS 2008, but also the Spanish CIS 2013 and the German BAS 2018. The Polish scale correlates with the MPS and our scales from 1990 and 1999 at only .27, .11, and .25, respectively. These values are statistically insignificant,²¹ and such low values are extremely rare.²² The correlation with the BAS is markedly higher at .57, but this still represents the second-weakest association with the prestige scales overall. The correlations of the Spanish scale with the MPS and our 1990 and 1999 scales are only .51, .58, and .65, respectively. These values are also unusually low. The correlation of the CIS with the BAS is .79, which, from the BAS's perspective, is the second-highest value (the agreement with the USA 2016 is only marginally higher at .80).

The MPS and our two scales exhibit very similar values, despite differing numbers of occupations considered. For the prestige scales, correlations range between .79 and .95, with only the comparisons to China 1994 and NORC 2010 averaging closer to .80 than to .90.

To understand the poor performance of the BAS, a comparison with the three marginal scales is helpful. The AI scale of boast-worthiness can certainly be viewed as an indicator of prestige. Here, the BAS shows its lowest correlation, so this aspect offers no explanation. The AI scale of training/pay conceptually aligns closely with socioeconomic status. Here, the MPS, 1990, 1999, and BAS show correlations of .59, .61, .67, and .68, respectively. For the social-value scale by Newlands and Lutz, the values are .60, .54, .57, and .64. This means that the two scales where the BAS exhibits slightly higher correlations relate to the aspects least associated with occupational prestige. It should also be noted that our scales correlate more strongly with the international scales than with the German ones, and this even holds for Counts' century-old survey.

Furthermore, it is worth mentioning that the AI-generated prestige hierarchy shows some of the highest levels of agreement. This is not surprising, however, as the artificial intelligence draws its knowledge from available publications.

To conclude, we compare our 1999 scale with the scales from China 1994, Poland^[20], Spain^[24], the USA 2016, Hungary 2016, Great Britain 2018, Sweden 2018, and South Korea^[29] at the occupational level. Here, too, we limit ourselves to cases where the absolute difference in z-scores is at least 0.75 (with our 1999 scale as the subtrahend).

Of the 37 common occupations with China 1994, 6 exceed the threshold. In China, the politician enjoys higher prestige than in Germany (1.50); in Germany, the programmer (-1.24), the master painter (-1.17), the goldsmith (-1.15), the physician (-1.14), and the librarian (-0.87) have relatively higher prestige.

With the Polish CBOS 2008, there are hardly any similarities. Fourteen of the 19 occupations differ by more than three-quarters of a standard deviation. The largest differences are found for the politician (-2.41), the firefighter (1.80), and the pastor (-1.64). In Poland, the firefighter ranks second, the pastor ranks second to last, and the politician ranks last.²³ We will return to these three occupations in detail.

A wealth of differences is also evident with the Spanish survey. Of the 54 common occupations, 19 exceed the threshold. Table 15 presents the z-score differences and the rankings.

		Rank	
Occupation	z-diff.	GER 1999	ESP 2013
Firefighter	2.11	23	1
Primary School Teacher	1.37	24	5
Bus Driver	1.22	47	29
Butcher	1.07	44	27
Auto Mechanic	0.99	37	20
Educator (Kindergarten)	0.95	38	21
Hairdresser	0.94	45	32
Social Pedagogue	0.91	32	12
Worker (Assembly Line)	0.77	55	47
Nurse	-0.78	35	50
Bank Employee	-0.78	26	40
Manager (Car Industry)	-0.82	6	19
Real Estate Agent	-0.83	31	44
Ambassador	-0.89	3	14
Interior Designer	-1.20	12	35
Judge	-1.23	2	18
Savings Bank Branch Manager	-1.26	20	41
Pastor	-1.90	22	53
Politician	-2.78	27	55

Table 15. Differences Germany 1999, Spain 2013. (*italics=red; bold=green*)

As in Poland, the largest differences by far are found for the politician,²⁴ the firefighter, and the pastor. In Spain, the firefighter ranks first, while in Germany he ranks 23rd. Conversely, the politician ranks last in Spain but 27th in Germany. The pastor ranks third to last (53rd) in Spain, compared to 22nd in Germany. Naturally, occupations that perform better in Spain tend to rank in the lower half in Germany, and vice versa.

With the British survey from 2018, there are also 54 common occupations, but, as shown in Table 16, only 7 exceed the threshold.

		Rank	
Occupation	<i>z</i> -diff.	GER 1999	GBR 2018
Firefighter	1.26	18	7
Insurance Agent	0.97	48	28
Real Estate Agent	<i>-0.75</i>	24	38
Pastor	<i>-0.82</i>	15	26
Journalist	<i>-0.87</i>	10	20
Bank Employee	<i>-0.95</i>	20	37
Traffic Policeman	<i>-1.25</i>	31	53

Table 16. Differences Germany 1999, Great Britain 2018. (*italics*=red; **bold**=green)

In Great Britain 2018, the firefighter also enjoys higher prestige and the pastor lower prestige compared to Germany 1999. Additionally, the insurance agent holds a relatively higher position in Great Britain, while the traffic policeman, bank employee, journalist, and real estate agent perform better in Germany.

In the comparison with the USA 2016, only 2 of the 49 common occupations exceed the threshold. The firefighter again performs relatively worse in Germany, but the *z*-score difference is only 0.82. For the bank employee (-1.12), the relationship is reversed.

In the comparison with Sweden 2018, only the firefighter (0.84) exceeds the threshold. Among the 34 common occupations, firefighters rank 18th in Germany and 11.5th in Sweden.

Of the 44 common occupations with Hungary 2018, the accountant and the insurance agent perform better than in Germany (1.14 and 0.82, respectively), while the photographer and interior designer are rated higher in Germany (-0.81 and -1.23, respectively).

The comparison with South Korea 2022 yields only one difference: The congressperson ranks first in South Korea, while the (only partially comparable) politician ranks seventh in Germany (z-scores 1.53 and 0.20).

Excluding the special cases of Poland and Spain, the differences between Germany 1999 and the other countries are minor, and no consistent pattern is discernible. The only exception is the firefighter, who fares better in the USA, Great Britain, Poland, Sweden, and Spain, but occupies nearly the same rank in Hungary (16 vs. 15).

An Extreme Group Comparison

Overall, occupational prestige ratings present a remarkably consistent picture. Nevertheless, there are some differences, and these exhibit a certain degree of systematicity. To shed light on these divergences, we conclude by examining two extreme groups from our main study in 1999. On one side are the 10 participants with the highest agreement with the group standard, and on the other side are the 10 participants with the lowest agreement. The index $r(i, \text{rest})$ ranges from .889 to .991 for the orthodox group and from .480 to .609 for the unorthodox group.

The correlation between the group standards is .741, meaning that the shared variance is three times greater than the unshared variance, even among these extreme groups.

Table 17 presents statistical metrics of the group standards.

	Unorthodox	Orthodox	Total
Mean	50.1	54.6	52.1
Standard deviation	12.3	17.6	18.1
Minimum	23.8	18.0	17.7
Maximum	78.4	88.3	91.7

Table 17. Orthodox versus unorthodox raters and overall group.

The orthodox participants, on average, assign higher ratings than the unorthodox participants (54.6 vs. 50.1); the overall group mean of 52.1 lies approximately equidistant between them. A fundamental difference is observed in differentiation. With values ranging from 18.0 to 88.3, the orthodox group shows roughly the same range as the overall group (17.7 to 91.7). In contrast, the unorthodox group's values range only from 23.8 to 78.4. This indicates that the unorthodox group exhibits strong leveling, with the upper range being more compressed than the lower range.

Table 18 presents the occupations for which the z-scores differ by at least three-quarters of a standard deviation. Positive values indicate that the occupation has relatively higher prestige among the orthodox group, while negative values point to a relatively higher rating by the unorthodox group.

Occupation	Total	Unorth.	Orthod.	zO - zU
Politician	55.7	32.2	74.0	2.56
Ministerial Councillor	70.8	42.5	76.4	1.86
Pastor	63.0	39.6	70.0	1.73
Opera Singer	67.4	44.1	72.4	1.50
Real Estate Agent	52.0	40.4	66.4	1.47
Psychologist	65.1	46.6	71.1	1.23
Insurance Agent	28.8	23.8	36.3	1.11
Student (Business Admin.)	40.9	37.0	55.0	1.09
Opinion Researcher	44.8	32.6	48.4	1.08
Concert Pianist	70.9	55.4	80.4	1.04
Sculptor	53.9	41.0	58.0	0.94
Translator	58.7	65.1	62.1	-0.79
Lifeguard	27.2	39.8	25.7	-0.80
Master Painter	38.4	46.7	35.4	-0.81
Pilot	83.8	75.7	76.0	-0.87
Commercial Artist	67.4	68.9	65.6	-0.91
Criminal Investigation Inspector	68.5	66.5	61.3	-0.95
Train Driver	41.5	57.8	48.3	-0.98
Firefighter	56.0	65.0	58.4	-0.99
Truck Driver	27.8	48.5	34.8	-0.99
Baker	43.0	56.4	46.0	-1.00
Retail Salesman	44.1	47.7	32.4	-1.06
Traffic Policeman	43.5	60.0	49.4	-1.10
Auto Mechanic	42.4	51.8	37.6	-1.10

Occupation	Total	Unorth.	Orthod.	zO - zU
Precision Mechanic	49.1	60.3	49.7	-1.11
Waiter	30.2	53.2	31.2	-1.58

Table 18. Differences between orthodox and unorthodox Raters. (*italics=red; bold=green*)

Eleven occupations have a higher relative position among orthodox raters compared to unorthodox raters. Of these, seven lie above the overall group mean. Here, the z-differences range from 0.94 to 2.56. Conversely, 15 occupations have a significantly higher relative position among unorthodox raters compared to orthodox raters, with nine of these below the mean. Here, the differences are notably smaller (-0.79 to -1.58). Overall, this also reflects the strong truncation of the upper range and the somewhat less pronounced truncation of the lower range by unorthodox raters. Given this pattern, it is not surprising that eight of the ten unorthodox raters are men, while eight of the ten orthodox raters are women.

Discussion

This work is titled „A Somewhat Different Perspective on Occupational Prestige Ratings.“ Over large parts, it follows traditional lines, but it illuminates and delves deeper into some aspects that are rarely or not at all considered. The first concerns individual response styles in terms of rating levels and differentiation between occupations. The second concerns the consensus among raters, which we examine at different levels. A third distinctive feature is the combined consideration of means and dispersion. Related to this, as a fourth point, is the consistent distinction between the absolute and relative position of occupations in the prestige hierarchy.

We did not begin presenting the results by examining occupations but rather with the response styles of the raters. Here, we demonstrated substantial differences in both rating levels and differentiation between occupations. Some participants predominantly assign ratings from the lower range, others show the opposite pattern, and still others primarily assign ratings from the middle range. Of particular interest is the dispersion. Some participants limit themselves to a very narrow range, while others utilize the scale's breadth to a much greater extent. The degree of differentiation gives raters different weights. Those who

use the full scale range have a greater influence than those who, for example, only assign values between 30 and 80.

Against the backdrop of these different response styles, one might assume it would be very difficult to establish a hierarchical order of occupations. The opposite is true. Regarding the relative position of occupations, there is a very high level of consensus. On the topic of consensus, we considered various aspects: the concordance between individual raters, the agreement between individuals and the group standard of the other participants, the reliability of the group standard, and the correlation between different group standards.

The fundamental building block is the agreement between individual raters. In our 1999 survey, this averages .594, which is nearly identical to the value of .60 reported by Treiman^[1]. This figure can be considered typical for occupational prestige ratings. Such a level of interindividual agreement is very rare in subjective judgments in other domains. This high level of consensus has several statistical implications.

The first implication concerns the agreement of individuals with the group standard of the other raters. This is necessarily much higher than interindividual concordance and, in our case, averages .779. The prestige hierarchy derived from the group is shared to a very high degree by most raters. The shared variance is at least 48 percent, and for 83.3 percent of participants, it exceeds 70 percent.

The second implication concerns the reliability of the group standard, measured with Cronbach's alpha. According to a common rule of thumb, alpha values in the .70s are considered acceptable, in the .80s good, and above .90 excellent. The reliability of our 1999 prestige scale is .995, practically perfect.

The third implication concerns the differentiation of subgroups. If individuals correlate on average at .60, it is impossible to form subgroups in a meaningful way that lack commonalities. Even in the comparison of the extreme groups—orthodox versus unorthodox—the commonalities are three times greater than the differences.

Implications for Study Design

The high agreement, which is greater for occupational prestige than for many other subjective judgments, has a significant implication for study design. If one is solely interested in a reliable ranking, a very small sample of raters is sufficient. This was also emphasized by Goldthorpe and Hope (^[48], p. 174): „assessments with a reliability in excess of 0.9 may be achieved by averaging gradings over no more than $k = 10$ respondents when the individual respondent has a reliability in the region of $r_{tt} = 0.6$.” Even if individual reliability is set somewhat lower, a prestige hierarchy with two dozen raters is extraordinarily reliable, and

with three dozen raters, one is certainly on the safe side. For example, if five samples of three dozen raters each evaluated five different lists of 100 occupations, with 180 participants, one would obtain a hierarchy for 500 occupations with a reliability above .96, likely even above .98.²⁵ To align the different lists to a common scale, as is occasionally practiced, one could use, for example, 30 occupations in each list. Then, with $7 \times 36 = 252$ participants, one would have a highly reliable ranking for $7 \times 70 + 30 = 520$ occupations. With 504 participants, one could cover 1010 occupations. Where the focus is solely on the prestige hierarchy, the very high rater agreement allows for significant progress with small rater samples, especially when limited to a manageable number of occupations. For example: We divided the 136 participants of our main study into four groups of 34 each, achieving alpha values of .975, .977, .982, and .982. The difference from .995 for the entire group is negligible. The correlation between the group standards ranges from .966 to .986. Since our participants took part voluntarily and without compensation in the online survey, the larger sample only required a slightly longer data collection period but incurred no costs or additional effort. Many studies, however, involved enormous effort, time, and costs. If the primary goal is to create a reliable hierarchy, significant savings can be achieved through the considerations outlined above. However, if additional questions are of interest, a much larger sample may be necessary.

Interpretation of Differences and Changes

A central aim of this work is a critical examination of the interpretation of differences and changes. The rule of thumb cited in the introduction by Treiman^[1], that differences of less than 6 points (on the theoretically 0-to-100 SIOPS scale) are not meaningful, has some merit, but focusing solely on mean differences falls short. Statistical significance and effect size arise only from the interplay of difference and dispersion, and beyond that, practical significance would also need to be assessed.

With the commonly used bounded category scales, it must be considered that mean and dispersion are not independent. An occupation can only be near the upper or lower bound under regular conditions if raters show high agreement. In the broad middle range, various combinations of mean and dispersion are possible.²⁶

This empirical relationship was pointed out by Counts^[7] a hundred years ago,²⁷ and other authors have addressed it without noting the mathematical necessity (e.g., ^{[8][33][49][34][30]}).

Moreover, it should be noted that prestige scores in the middle range are more densely clustered than at the extremes. This is illustrated in Figure 10 through the frequency distribution of our 1999 scale.

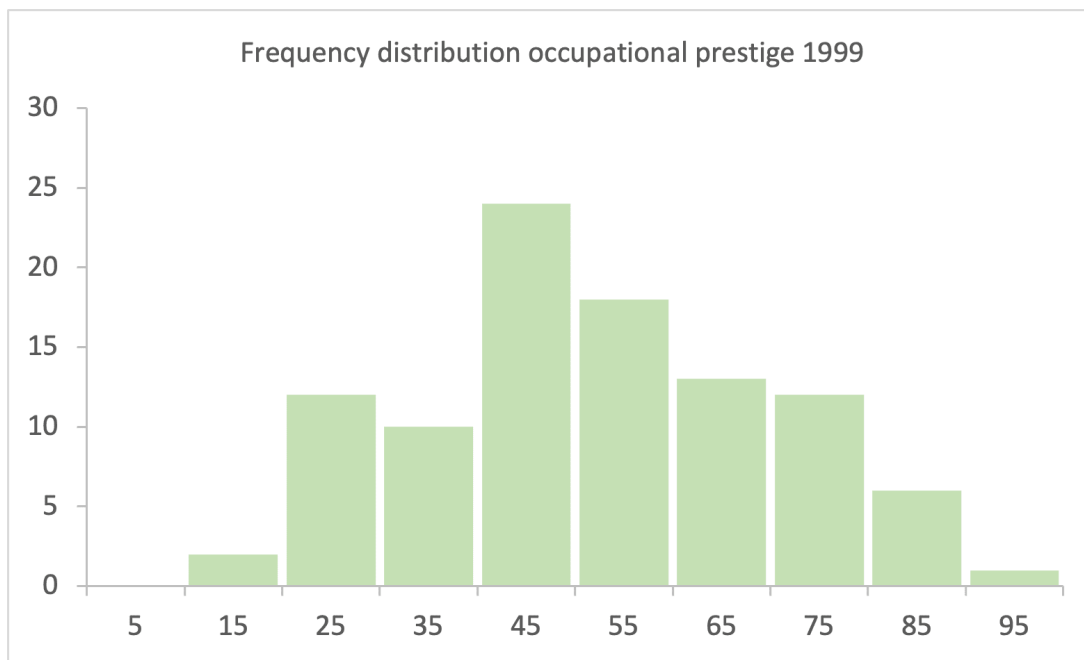


Figure 10. Frequency distribution occupational prestige 1999.

All this leads to the fact that absolute differences and rank differences can, in some cases, convey very different impressions. We have demonstrated this with several examples. It is worth recalling here that a gain of a few points can lead to a significant loss in rank order. Media, professional associations, and other institutions may highlight one perspective or the other depending on their interests, but science should raise awareness of this issue and always report means and dispersions, weighing absolute and relative differences against each other.

Occupational Prestige as a Psycho-Sociological Universal

The fundamental insight of a century of occupational prestige research is:

Regardless of the sample considered, the relative order of occupations is, *in its essentials*, nearly always and everywhere the same.

The consistency of prestige ratings across space and time is sometimes referred to as the Treiman Constant. Treiman^[1] describes the prestige hierarchy as „a genuine Durkheimian social fact,” and Hout and DiPrete^[50] consider it „the only universal sociologists have ever discovered—not just in stratification, but in sociology as a whole.”

Our analyses provide further confirmation. The extraordinarily high consistency is found within our surveys across different subgroups, between our surveys conducted 10 years apart, in comparison with German studies since the 1950s, and in comparison with international studies spanning a century.

Within our own surveys, we compared various group standards and obtained correlations around .97. Even the extreme comparison of orthodox versus unorthodox, conducted solely for demonstration, yields a correlation of .74.

The intra-German comparisons, with the exception of the BAS, show values between .84 and .90. Compared to Treiman (^[1], p. 43), who found correlations above .95 in 24 of 31 intra-national comparisons, our values are somewhat weaker. The most recent and comprehensive German study, the BAS, performs even worse. The correlation with our 1999 and 1990 scales is .76 and .74, respectively, which is on the order of our extreme orthodox versus unorthodox comparison. The correlation between the MPS and the BAS, at .62, is entirely an outlier^[38]. Such low values are rarely found in the literature.

The special role of the BAS is also evident in international comparisons. Here, the correlations are mostly even lower, in some cases below .70. In contrast, our 1990 and 1999 scales, as well as the MPS, perform significantly better in international comparisons than in intra-German ones, with values mostly between .90 and .95. Strikingly low is the agreement with Spain 2013, and even more so with Poland 2008, which is entirely an outlier.

At this point, the SIOPS, based on surveys from the 1950s and 1960s, should be highlighted. The correlation with our 1990 and 1999 scales is .88 and .84, respectively. Somewhat lower values are reported, for example, by Ulfssdotter-Eriksson and Nordlander for Sweden 2002 and 2018 (.83, .82). There is much to suggest that the classic scale, considered by some to be the gold standard, is no longer entirely up to date. Developing an updated international scale following the SIOPS model calls for collaboration within an international research community.

Despite some exceptions, there is no doubt that the prestige hierarchy, in its essentials, remains the same across space and time, although some systematic differences and trends can certainly be identified.

Variations and Trends: Sex and East-West Germany

When it comes to searching for differences, the most obvious variable is sex. The occupational domains of men and women have always been very different, and despite the increasing participation of women in the labor market, their experiences in the professional world remain distinct. However, it must be considered that this is not about personal experiences or preferences. It is about opinions regarding public opinion,

which consists equally of men and women. If participants follow the instructions and both sexes perceive public opinion similarly, no differences are to be expected. Indeed, the group standards of male and female raters show extraordinarily high agreement (.972). Nevertheless, there are systematic differences, and these manifest at various levels.

Within both sexes, there is very high interindividual concordance, but agreement is slightly higher among women than among men (.627 vs. .564).

In assigning scores, men and women agree in the lower quartile of the value range, but at the upper end, women more frequently assign very high ratings. This is particularly true for the maximum value of 100. Overall, women assign higher average scores than men.

At the occupational level, women show greater variability. Depending on the perspective, this can be described as greater differentiation by women or greater leveling by men.

Women rate top positions much more positively than men, while men rate lower occupations slightly higher than women. Overall, there is a remarkably strong relationship between occupational prestige and the difference between men and women. For absolute values, the correlation is $-.521$, and for effect size, it is even $-.635$.

From the perspective of an evolutionarily grounded mate value theory, this pattern appears plausible. A man's contribution to joint reproductive success lies primarily in protection and material provision, making social status a crucial, if not the most important, component of a man's mate value from a woman's perspective^{[51][42][43]}. Thus, it is understandable that women differentiate more strongly between occupations and grant a premium to top occupations, resulting in higher concordance.

Despite its plausibility, caution is warranted. There are studies that support one or another point, such as Lin and Xi^[52], Ulfsdotter Eriksson^{[53][54]}, and Valentino^[55], but a systematic analysis is not known to us, and it remains an open question to what extent our findings can be generalized. Future studies should routinely consider these aspects, and it would be desirable to evaluate existing datasets in this regard.

The sex variable has another aspect, namely the differentiation between occupational holders. Since the occupational domains of men and women differ greatly, one might assume that some occupations are rated differently depending on whether they are performed by men or women. However, here too, the prestige hierarchy is almost identical. For the 94 occupations in 1990 and the 34 occupations in 1997, the correlation is $.984$ and $.970$, respectively. Beneath this very high agreement, there is a weak tendency for some occupations to be rated more favorably for the sex that is more strongly represented among the

occupational holders. However, this applies to very few occupations, and the differences are mostly negligible. For example, among the 94 occupations in 1990 and the 34 in 1997, only two differences in each case exceed 10 points. Dramatic differences, however, are reported by Jacobs and Powell^[56]. For 45 of 56 occupations, the difference is significant ($p < .01$). In 16 cases, men have an advantage ranging from 10.3 to 35.3 points, and in 21 cases, women are rated higher by 10.4 to 41.8 points. There is an extraordinarily strong relationship with the sex ratio: the correlation between the difference (men minus women) and the proportion of men is .92. Krueger, Ebner, and Rohrbach-Schmidt^[57] report for the BAS that occupations with stronger sex segregation have higher prestige than those with a balanced sex ratio. However, this only holds after statistically controlling for income and required educational level; otherwise, the correlation between prestige and the proportion of women is slightly negative (-.163).

At this point, a methodological note is warranted. The BAS instructions include the passage: „To keep it brief, we almost consistently use the masculine plural form for occupational titles, e.g., ‘Ärzte’ instead of ‘Ärztinnen und Ärzte.’ However, please always consider both sexes”^[38]. Under this condition, it is unclear which sex raters have in mind, and some interesting questions cannot be explored at all. When the sex of occupational holders is the focus, it is necessary to explicitly differentiate between male and female occupational titles and specify which sex the rating should refer to. Overall, the evidence on this topic is quite sparse and inconsistent^[38].

Due to its unique historical development, a comparison between East and West Germany is always relevant. Here, too, there is very high agreement between group standards (.961), and the average rating is practically the same in the West and East (51.9 vs. 52.2). Nevertheless, we observe some differences. First, West Germans show greater differentiation, or conversely, East Germans exhibit stronger leveling. Second, occupations rated significantly higher in the West than in the East are found in the upper range; occupations rated significantly higher in the East are found in the lower range; and across all occupations, the correlation between prestige and the West-East difference is .591. This reveals an interesting parallel to the sex comparison: the West German pattern corresponds to that of female raters, and the East German pattern to that of male raters. In the East-West comparison, it must be considered that our survey was conducted in 1999, a decade after the fall of the Berlin Wall and 9 years after reunification, and that the East German sample consists of only 13 participants (thanks to the very high rater agreement and balanced sex ratio, the comparison has a reliable basis). Since the East-West divide is still evident in many areas today—consider, for example, the results of the 2025 federal election, which reveal a radical division exactly along the former border—an updated comprehensive assessment would be desirable.

The considerations within the framework of our own surveys have revealed, beneath the surface of overwhelming agreement, some smaller but substantively plausible differences. These, however, require further empirical confirmation or an update after a quarter-century of significant changes in the professional world and cultural values.

In comparisons with other studies, it is worth emphasizing once again that our prestige hierarchy is more closely aligned with international surveys than with German ones. This applies in a negative sense particularly to the Berufliche Ansehensskala (BAS) and in a positive sense to the more recent studies from the USA, South Korea, and Sweden. Comparisons with the latter have also revealed some minor differences, but these do not form an interpretable pattern. In comparisons with international studies, the issue of occupational matching is always a challenge. Here, we deliberately limited ourselves to relatively clear cases, accepting a correspondingly smaller sample of occupations. Since different occupational groups and the full spectrum are represented, there is no reason to assume that the relationships are significantly over- or underestimated.

Decliners, Climbers, Fixed Points

In our previous comparisons, we limited ourselves to occupations common to our list; now, we consider the entirety of occupations. To ensure comparability despite different numbers of occupations and scale usage, we normalize the prestige scores and dispersion to the range of 0 to 100.²⁸ Where it seems more illustrative, we occasionally provide rank positions. As exemplary cases, we consider the pastor, the politician, the physician, and the firefighter.

A hundred years ago, in the USA, the pastor ranked 4th among 45 occupations, achieving a value of 91. In the international SIOPS, constructed from studies in the 1950s and 1960s, the pastor scores 86. In the Spanish CIS 2013 and the German BAS 2018, with values of 31 and 23, the pastor is in the lower third. However, as emphasized, the CIS and BAS are outliers, yet a clear trend indicates a significant loss of prestige. In the most recent surveys considered in this work from Germany, Great Britain, Poland, Sweden, Hungary, and the USA, the pastor is only in the middle range. Beyond the position in the ranking, another point is of particular interest, namely dispersion. As mentioned in the introduction, Treiman^[1] noted that among 50 common occupations, opinions about the pastor diverged the most; Bolte^[33] reported the same for his Schleswig-Holstein sample. In our 1999 survey, the pastor has the second-highest dispersion. In the BAS 2018, the pastor scores 91, in Spain 2013, 82, and in Great Britain 2022, 74.

The situation is similar for the politician. In Smith^[8], politicians of the four highest ranks in the USA followed Supreme Court judges at the top. Treiman^[1] also reports that various high-ranking political offices are at the top, but since they are not occupational titles, they are not included in the SIOPS. For Germany, we can paint a fairly broad picture. In Wurzbacher^[58] and Bolte^[33], the government councillor still ranks high, and in Kleining and Moore^[34], the envoy in diplomatic service, the lord mayor, the ministerial councillor, and the district administrator achieve values of 94, 87, 87, and 80. In our 1990 survey, the politician scores 67, but in 1999, only 51. In the BAS 2018, the mayor is still in the upper half with 60, but the state parliament member and federal parliament member score 46 and 39, in the lower half. Extreme cases are Spain 2013 and Poland. In Spain 2013, members of the central government and national congress rank second to last and third to last, respectively, while city councillors and regional parliamentarians score only 14 and 19; however, the ambassador enjoys high prestige (78). In Poland, the minister, city council member, parliament member, and party official occupy (in descending order) the four lowest ranks. The opposite extreme to Spain and Poland is South Korea 2022 and China 1994, where the politician ranks 1st and 3rd, respectively. In Great Britain 2022, Sweden 2002 and 2018, and Hungary 2013, politicians also enjoy high prestige. As with the pastor, dispersion is of particular interest for the politician. In Bolte^[33], the government councillor is still rated very consistently, and for the various politician groups in Kleining and Moore^[34], dispersion is in the middle range. In contrast, in our 1999 survey, the politician has the highest dispersion. The same applies to China in the 1980s²⁹ and South Korea 2022. Extremely large disagreements are also found in Sweden 2002,³⁰ Spain 2013, and Great Britain 2022. Our 1999 survey also shows a remarkable result mentioned in the results section: women rate the politician much more favorably than men (62 vs. 36). Thus, the very high dispersion for the politician is partly a sex effect. The sex difference fits the evolutionary psychological perspective mentioned earlier, but it is unclear to what extent this finding can be generalized.

Thirdly, we consider the physician, and here we can be brief. Physicians of various specialties consistently rank very high across space and time, often in 1st place, and dispersion is usually very low. A curiosity is worth noting: in the BAS, raters are extraordinarily consistent for chief physicians, dentists, and veterinarians, while for general practitioners, dispersion is in the middle range, but physicians have the seventh-highest dispersion among 402 occupations.

Finally, we consider the firefighter. In the international SIOPS 1970, the firefighter scored 39. In China 1994, 43. In our surveys, the firefighter improved from 36 in 1990 to 52 in 1999. In the NORC, the firefighter improved from 52 in 1989 to 69 in 2012,³¹ in Spain from 78 in 1991 to 95 in 2013,³² and in Sweden from 64

in 2002 to 77 in 2018. In Poland 2008, Hungary 2013, Great Britain 2022, and the German BAS 2018, the firefighter achieves values of 97, 84, 89, and 89, respectively.

To explain the firefighter's meteoric rise, as well as the ratings of the pastor, politician, and physician, it is worth examining a study by the GfK Verein^[5]. In autumn 2017, 23,000 participants in 20 countries were not asked about occupational prestige but about the trust they place in 32 occupations. The firefighter ranks first in Brazil, Germany, France, Iran, Italy, Poland, Russia, Sweden, Spain, South Korea, and the USA. In India, South Africa, and Turkey, the physician ranks first. In Great Britain, Japan, and the Netherlands, the paramedic ranks first. In Indonesia, the teacher. In Mexico, the nurse. In Kenya, the farmer. In 17 countries, the politician ranks last, in Iran and Indonesia, the pastor, and in Sweden, the salesman. Across all countries, the firefighter, physician, paramedic, and nurse rank at the top. The pastor ranks 26th, in the lower quartile. The politician is by far the least trusted, with only 29 percent approval, the only occupation trusted by less than half of respondents. In the Global Trustworthiness Index 2022^[59],³³ the physician ranks at the top everywhere, mostly in 1st place, the pastor in the lower half, the government minister second to last, and the politician (politicians generally) in last place.

Polarization? No, Leveling

With the pastor, politician, physician, and firefighter, we have focused on four specific occupations; now, we consider the changes from a different perspective. Social development in Western countries is characterized by increasing polarization in many areas, and given the fundamental changes in the professional world, the question arises whether a corresponding development is also evident in occupational prestige^{[53][25][60]}. Two aspects can be distinguished here. First, whether the gap between top and bottom has widened over time, and second, whether raters have given more extreme ratings over time. In both cases, the answer is a clear no. In our 1990 survey, prestige values ranged from 6 to 92, in 1999 only from 18 to 92. In the BAS, the range in 2018 was even narrower, from 31 to 87, despite covering four times as many occupations. In Sweden, from 2002 to 2018, the value range narrowed from 8 to 92 to 14 to 89, and in Poland, from 1975 to 2008, the interval shrank from 32 to 90 to 42 to 81. In all cases, the gap between top and bottom has noticeably decreased, with lower occupations gaining more than higher ones have lost. A corresponding development in the NORC was cited in the introduction: „lower status occupational titles did gain disproportionally from the upgrading”^[14]. From NORC 1989 to NORC 2012, the maximum dropped from 86 to 78, thus reducing the gap from the other side as well. In all these cases, we see not polarization but rather a leveling of differences. An extreme case of leveling in individual scale

usage was cited in the introduction for Poland: „More and more respondents ... restrict the rating scale to a small range, while fewer and fewer are willing to differentiate prestige to the extent the scale allows”^[20]. In the BAS’s 0-to-10 scale, less than 1 percent of ratings are 0, and less than 3 percent are 10. In our 0-to-100 scale, 3.2 percent are below 10, and 4.6 percent are above 90. From this perspective, too, there is no talk of polarization.

The convergence of top and bottom pertains to the absolute rating of occupations. The crucial point is that the relative order has largely remained the same. Stability and change are not incompatible opposites. Even with extremely high correlations over time, a few occupations may exhibit significant changes, and the degree of differentiation between occupations can shift. The same applies to the comparison of group standards in cross-sectional analyses. Depending on interests, one may focus on commonalities or differences, but the other side should not be overlooked.

Treiman Constant: Stability and Change

Empirical research on occupational prestige celebrates its centennial this year, thanks to George S. Counts. Nearly half a century ago, Donald J. Treiman provided a work that remains unmatched in empirical and theoretical terms. On the one hand, he synthesized empirical findings from over 50 countries at different developmental stages worldwide into a common prestige hierarchy, the SIOPS. On the other hand, he convincingly argued that this order necessarily arises from the division of labor, which follows the same logical principles in all human societies, and that the resulting inequalities are judged very similarly by people worldwide. We are thus dealing with an interplay of sociological and psychological regularities, roughly summarized in the Treiman Constant: The prestige hierarchy of occupations is, *in its essentials*, the same across space and time.

While empirical and theoretical objections have been raised, all attempts to dismantle the Treiman Constant in its essentials have failed. By comparing our 1990 and 1999 prestige hierarchies with surveys from China, Germany, Great Britain, Poland, Sweden, Spain, South Korea, Hungary, and the USA from 1925 to 2024, we have provided further evidence for the validity of the Treiman Constant across space and time, but also clear evidence of some spatial and temporal differences.

In the following, we focus on the differences, as these contribute more to theoretical understanding than the commonalities. In Germany, this concerns the BAS from 2018, in Spain the CIS^[24], and in Poland the CBOS 2008. These three deviate significantly to radically from the usual pattern.

We begin with the GfK survey from 2017, which did not address occupational prestige but rather the trust placed in 32 occupations across 20 countries. The correlation of global GfK trustworthiness with occupational prestige in our main 1999 study is .155, which is statistically insignificant ($p = .492$, $df = 22$). Our prestige hierarchy has nothing to do with trust. The situation is entirely different for the BAS, CIS, and CBOS. Here, the correlations are .497, .733, and .935, respectively. Thus, the shared variance between prestige and trust is 25 percent for the BAS, 54 percent for Spain, and 87 percent for Poland. For the BAS, the relationship is statistically significant ($p = .008$, $df = 27$), but a partial overlap of prestige and trust is not problematic, as it makes sense that the valuation of occupations is influenced, among other factors, by trustworthiness. In Spain, however, the overlap is concerning. The CIS^[24] correlates more strongly with GfK trust than with our prestige scales and the MPS. The extreme case is Poland, where prestige and trust are nearly identical.³⁴

Now we recall the politician, pastor, and firefighter, which have seen spectacular changes in prestige. In Spain and Poland, politicians have plummeted from the top to the lowest ranks. In the BAS, the federal and state parliament members are in the upper part of the lower half. In all cases, the pastor is in the lower third, and the firefighter is found at the very top. These ratings align with the general trend, but in Spain and Poland, we see them in their most extreme form, and in the BAS in a moderated form.

The three „problem cases” are characterized by an unusually strong leveling of differences between occupations, particularly affecting the lower range. In our 1999 scale, 17 percent of occupations are in the lower third of the value range; in 1990, it was even 23 percent. In contrast, in Spain, only one percent are in this range, in the BAS only half a percent, and in Poland, none at all.

The Polish CBOS 2008 is an anomaly in several respects. It shows no correlation with our scales and the MPS, measuring a different construct—trust—instead of prestige. The collapse of various politician groups is in line with the trend but goes far beyond the usual extent. The same applies to the almost complete absence of the lower scale half. The extreme leveling is certainly a key reason why the interrater correlation is much lower than typical for occupational prestige ratings. However, this does not explain why this pattern is found in Poland. Domański et al.^[20] attribute their findings to general uncertainty following the radical political, economic, and social upheavals after the collapse of socialist rule. It should be noted, however, that Poland already showed remarkably low agreement with other countries in the 1950s.³⁵

The Spanish CIS^[24] is a borderline case. Judgments are strongly influenced by trust, and several changes from 1991 to 2013, not only those of the politician, pastor, and firefighter, are extreme, but they follow a

widespread trend.

The situation with the BAS is not as severe, but it also warrants scrutiny. Given its moderate agreement with our own and international scales,³⁶ it is unclear what reflects genuine temporal changes, specifically German peculiarities, or methodological weaknesses. The fact that participants rated only 5 occupations each is highly problematic. This extreme reduction is justified by the collection of additional variables. However, with only 5 out of 402 occupations, there is a risk of context effects.³⁷ The minimal sets inevitably cover different segments, providing participants with different reference systems. The same issue, to a lesser extent, arises with the Spanish CIS, where the 285 occupations were divided into 19 sets of 15 each. Our own analyses show some non-trivial differences between segments. For the BAS, this must be even more pronounced. Before elevating the BAS to the new standard for the German occupational prestige hierarchy, a pure prestige study with much larger sets should be conducted.³⁸ For example, the 402 occupations could be divided into 4 sets of 100 or 101, covering the spectrum equally. With a typical interrater correlation, $4 \times 50 = 200$ participants would yield a reliability above .98, and even with a sample of 100, a reliability above .96 could be expected.³⁹ For comparison: the BAS had 9,011 participants. To enable greater differentiation, a 0-to-100 scale should be used instead of a 0-to-10 scale. Additionally, sex should be clearly and unambiguously specified. If the sex of occupational holders is of specific interest, two separate lists should be used.

Regarding interrater correlation, we would like to address a point not yet discussed. When dividing a sample into different subgroups—e.g., men and women, old and young, rich and poor, educated and uneducated—the group standards typically agree to a very high degree. This independence from group composition is the empirical basis of the Treiman Constant. When conducting group comparisons at a different level, a different picture emerges. There are good theoretical reasons and well-founded empirical evidence that the degree of rater agreement depends on the group considered. Within some groups, concordance is very high; in others, perceptions of the professional world and public opinion are more diffuse, and consensus is lower. A classic example is the study by Guppy^[61]. His starting point was the studies by Hodge, Siegel, and Rossi^[10] in the USA and Pineo and Porter^[62] in Canada. He compared majority with minority groups (Whites/Blacks in the USA and Anglophones/Francophones in Canada), white-collar with blue-collar occupations, and university education with high school and elementary school. The minimum correlation between group standards is .90 (university/elementary), with other values ranging from .96 to .99. Additionally, Guppy calculated the average correlation between different pairs at the individual level. In both countries, pairs from the majority group show higher consensus than

mixed majority-minority pairs, which in turn are more consistent than minority group pairs. The same pattern exists for white-collar, mixed, and blue-collar pairs, as well as for educational levels, with a descending order. Systematic differences are found in all areas, particularly pronounced in education, ranging in the USA from .34 (elementary/elementary) to .80 (university/university) and in Canada from .45 (elementary/elementary) to .84 (university/high).

Guppy appears to see a contradiction between his findings and the Treiman Constant, but that would be an error. Concordance within different groups pertains to a different level than the agreement of group standards. However, both are interconnected. The crucial point is: nearly all individuals share the general standard to a substantial degree. Technically speaking, the discrimination is positive for nearly all raters, and for most, it is very high. In our case, it averages .779 (Table 3), with a minimum of .480. Zero correlations are very rare, and where substantial negative correlations are observed, they are likely due to reversed polarity or a disruptive participant. If they were genuine reversals, the intermediate range would be much more populated. Individual commonalities with the general standard are accentuated through averaging, resulting in very high correlations at the group level, regardless of group composition.

From Guppy's perspective, social privileges are the common factor underlying his groupings; from a psychological perspective, another variable is obvious: cognitive abilities, which are ultimately a central cause of social privileges^{[63][64][65][66][67]}. All three of Guppy's distinctions involve educational level and intelligence.⁴⁰ In all cases, consensus among raters with higher education is stronger than among those with less education. This finding is immediately intuitive. More educated individuals also have more comprehensive knowledge of the complexity of the professional world and public opinion, and the typically longer shared education could lead to greater alignment of opinions. Individuals with lower cognitive abilities have a more diffuse worldview, and their knowledge is more closely tied to their own experiences. Of course, intelligence and education are not the only causes of group differences in rater agreement. For example, the (slightly) higher consensus among female participants in our studies is certainly due to other factors. Differences in consensus within different groups are an insightful topic that is insufficiently researched and deserves far greater attention. Reporting interrater correlations, ideally broken down by different rater groups, should become an essential component of research reports.

Consensus among raters is also evident from another perspective: the dispersion of different occupations. A hundred years ago, when statistics was still manual labor, Counts used the interquartile range of occupations to determine that consensus among teachers was greater than among students.⁴¹ On the same basis, Bolte concluded in the 1950s: „The clarity of classification thus increases with higher education and

age. Both factors apparently lead to a clearer crystallization of the perception of a prestige hierarchy”^[33]. The greatest disagreement was found among young workers, who included a high proportion of vocational school students, and in this group, the pastor was the occupation with the highest dispersion. Overall, the greatest agreement was at the top, with the professor and physician, and at the bottom, with the messenger and unskilled worker. At that time, the government councillor enjoyed the fourth-highest prestige, and raters were highly consistent in this regard. The firefighter was not considered at the time.

The pastor, politician, and physician, or their equivalents at the time, have enjoyed the highest prestige for centuries. However, their development in recent times has diverged.

For the physician, practically nothing has changed; due to the very high professional qualification requirements and the paramount importance of health, the physician continues to rank at the top almost everywhere. In their orbit, some medical and caregiving occupations have gained prestige, and the firefighter’s unprecedented rise can be situated in this context. This reflects, on the one hand, the growing importance of health and personal safety, and on the other hand, a tendency to conflate the concept of prestige with other factors such as social utility or trust, as well as to present oneself in a socially desirable light. We will return to this point.

A very different picture emerges for the pastor and politician. First, it must be noted that there are significant differences, especially for the politician. In South Korea, the congressperson still ranks first in 2022, while in the outliers Spain and Poland, various politician groups have plummeted from the top to the lowest ranks. In the other countries we considered, the pastor and politician have lost prestige to varying degrees. The fact that cultural differences and temporal shifts can be particularly evident in (originally) highly regarded occupations in the religious and political spheres, while otherwise high agreement and stability prevail, was recognized early on.⁴² The dominant position of religious and political leaders, once considered a natural order, is increasingly questioned in the context of a more rational worldview. The ongoing reevaluation is evident in a point we have repeatedly emphasized: for the pastor and politician, opinions diverge extraordinarily. For these two, we see both significant spatial and temporal differences between groups and particularly strong differences within groups.

For the vast majority of occupations, we see only minor differences across space and time, and among raters, we observe a very high degree of agreement, which is rarely found in subjective judgments. The Treiman Constant still holds, even though the prestige of no single occupation is fixed. The picture painted by Counts a hundred years ago would not surprise anyone today, although few would place the banker in first place, and only a few would rank the pastor fourth.

Leveling and the Transformation of the Prestige Concept? A Cognitive Perspective

In conclusion, we aim to clarify a specific methodological point and, in this context, highlight a topic that may seem marginal to some but which we consider highly significant.

We have repeatedly emphasized that considering means alone is insufficient and that dispersion must always be taken into account. In particular, we noted that occupations can only be near the upper or lower end of the value range if dispersion is very low. However, we also repeatedly mentioned the congressperson in South Korea, who both ranks first and has the highest dispersion. This is neither a contradiction to our point nor a South Korean curiosity. The scale by Kye and Seol ranges from 0 to 10. The congressperson has a mean of 8.18, which is far from the upper end of the scale. Similarly, the simple laborer with 1.83 is far from the lower end. Moreover, dispersion varies only from 1.60 to 2.10.⁴³ As a result, the relationship here is not rainbow-shaped but even slightly U-shaped.

This brings us to a topic we have touched on several times but consider more significant than it may initially appear. The peculiar situation in South Korea is due to the fact that the highest prestige value is far from the upper end and the lowest far from the lower end of the value range. This is likely partly because raters disagree and neutralize each other. The more substantial reason is probably that raters avoid the extreme ranges.

In this context, it is worth examining the BAS again. „The edges of the distribution, i.e., the prestige values 0 and 10, are not strongly represented percentage-wise, so that, in principle, no floor effects and, if any, only very minor ceiling effects are visible. From this perspective, the 0-to-10 rating scale seems to offer sufficient differentiation in assessing occupational prestige”^[38]. We see it exactly the opposite way. There is no talk of ceiling or floor effects, and by avoiding the upper and especially the lower category, differentiation is not increased but restricted.

The extreme case was seen in Poland. We provided specific figures in the introduction (p. 16), and we quote the conclusion here for the third time: „More and more respondents do not differentiate occupations by prestige at all, increasingly restrict the rating scale to a small range, while fewer and fewer are willing to differentiate prestige to the extent the scale allows”^[20]. We have already mentioned the consequences of this leveling, but we repeat them due to their importance. First, an extraordinarily low interrater agreement for prestige research. Second, an extraordinarily low agreement with other countries. Third, no occupation is found in the lower 40 percent of the value range. Fourth, instead of prestige, trustworthiness was assessed.

The Polish CBOS 2008 is an extreme case that is almost worthless for occupational prestige research but, in its extremity, points to a core issue that has been observed multiple times but, in our opinion, receives too little attention.

Leveling, as extensively described, is a widespread trend, particularly affecting the lower range. We view the upgrading of lower ranks as a consequence of moral advancement, which goes hand in hand with the cognitive advancement described by the Flynn Effect.

At this point, we must insert a digression. Psychometric intelligence tests were developed at the beginning of the 20th century and normed so that the average IQ is 100. In later tests, it was found that average performance had increased, requiring more difficult tasks to maintain the mean of 100. Consequently, IQ tests had to be repeatedly adjusted. In the Western world, IQ rose by 30 points in the 20th century, equivalent to 2 standard deviations, meaning that performances that were average in the original tests would fall in the bottom 5 percent in today's tests. This phenomenon, known as the Flynn Effect, is observed globally, albeit with varying gradients^{[68][69]}, and many countries today are where the Western world was a hundred years ago. It is impossible for this massive increase in phenotypic intelligence to have genetic causes.⁴⁴ What has changed are the fundamental structures of thinking. The genetic potential was present in earlier times,⁴⁵ but it was only activated through the massively increased societal complexity and the integration of the entire society into a more extensive and effective education system.⁴⁶ Intelligence encompasses a broad spectrum of cognitive abilities, and the Flynn Effect is evident only in some areas, while performance in others has remained more or less constant. What is new is the ability for formal-operational thinking in the sense of Jean Piaget's developmental psychology.⁴⁷ The ability for logical thinking has existed for thousands of years, but it was limited to the framework of one's concrete experiential world.⁴⁸ Formal-operational thinking is abstract, hypothetical thinking, independent of the concrete experiential world, capable of drawing logical conclusions about non-existent constructs. This necessary prerequisite for true scientific thinking was met in antiquity only by a small number of outstanding thinkers, and even then, only in some domains. From the 16th century, the first thinkers of this new type emerged in northwestern Europe, gradually increasing in number, but it was only in the 20th century that the ability for formal-operational thinking became widespread in the Western world and a few other countries. Even here, this ability is developed only in some domains for most people, and in everyday life, concrete-operational thinking remains the modal variant. In many countries with lower national IQs, the ability for formal-operational thinking is developed only in a very small portion of the population, and even then, only weakly^{[64][70]}.

What does this digression have to do with our topic? Cognitive ability determines how we perceive the world and how we judge our fellow humans and social structures. It determines the degree to which we can empathize with others, and it goes hand in hand with moral development^{[71][72]}. People who do not progress beyond the concrete-operational stage view social structures as natural givens. Top is top, and bottom is bottom. They consider, for example, representatives of secular and religious authority, such as the politician and pastor, as unquestionable authorities deserving obedience and recognition. Empathy is only weakly developed, and it seems self-evident that the upper ranks keep the lower ones down and oppress them. With formal thinking, the picture changes. Social structures are no longer seen as divinely ordained natural orders, and social rank is judged more rationally. Empathy increases, the radical gap between top and bottom is seen as unjust, and the lower ranks are elevated.

This roughly sketched picture is undoubtedly a positive development for people at the formal-operational stage. However, it poses a problem for occupational prestige research. The prestige of occupations follows a clear hierarchical order. We ask participants how they believe this order is perceived by the public. To this end, we provide scales of 1 to 5, 1 to 9, 0 to 10, or 0 to 100. The key point is: the endpoints are meant only to mark the lowest and highest prestige. We do not mean that very low prestige is associated with worthlessness or moral condemnation, nor do we mean that very high prestige implies godlike reverence. The increasing tendency to avoid the extreme ranges, particularly the lower one, suggests that a growing number of raters incorporate moral considerations, seek to establish justice in their personal sense, and wish to present themselves in a positive light. The conflation of „is” and „ought” is as undesirable as a strong conflation with other criteria such as social utility or trustworthiness. Such criteria undoubtedly influence prestige, but if they are weighted too heavily, they distort the picture. In our view, this is the case, for example, with the firefighter and some medical occupations with lower requirements. Occupational prestige, in our sense, is the esteem, respect, appreciation, or deference associated with occupations, which divides society into groups one looks up to, considers equal, or looks down upon. From this perspective, the firefighter definitely does not belong in the top ranks. Similarly, politicians are not at the very bottom of public esteem, even if we regard some of their representatives with disdain or contempt.

The preceding discussion concerns two distinct, though not independent, points: differentiation between occupations and the distinction of occupational prestige from other criteria. The ability to differentiate is a cornerstone of intelligence. Differentiation leads to varied evaluations and corresponding behavior; otherwise, it would be meaningless. The differing prestige of occupations is a social fact, and leveling for ostensibly moral reasons or evaluations based on personal ideals is, in our view, undesirable. Regarding the

distinction of criteria, the situation is mixed. On the one hand, prestige should not be conflated with conceptually and empirically distinct criteria such as social utility or trustworthiness. On the other hand, it is possible that such criteria are weighted more heavily by the public, leading to a substantial transformation of the prestige concept. The desirability of such a development is a matter on which opinions differ.

Notes

- An Excel file of the raw data for the main 1999 study is available at <https://osf.io/h9u8m>
- The German version of this work is available in Henss^[73].

Footnotes

¹ Wikipedia^[74] lists the origin of the surname Bergmann as „residential name (living on a mountain)”. However, Bergmann directly denotes the profession of miner. Source: https://de.wikipedia.org/wiki/Liste_der_h%C3%A4ufigsten_Familiennamen_in_Deutschland [2024/09/01].

² The Statista table is titled „Which of these occupations, in your opinion, enjoys high or no high prestige? (Share of respondents who assign very high or high prestige),” but a note on the side states: „The source does not provide precise information about the question wording,” making it unclear which criterion underlies this peculiar ranking.

³ We refer to the group standard as the values of the various occupations averaged across all members of the group.

⁴ In positions 96 to 100 are unskilled worker for odd jobs, scrub-woman, garbage collector, unskilled migratory worker, and professional prostitute.

⁵ The correlation of NORC 1963 with Counts^[7] and Smith^[8] is .934 and .971, respectively.

⁶ In the earlier NORC studies, the five categories were assigned the values 20, 40, 60, 80, and 100. Thus, the value range did not start at 0.

⁷ The coefficients for the post-socialist period are based on our own calculations.

⁸ How much do you have to learn to enter into this occupation? How much power or influence does it bring? How much money can be earned in this occupation? How useful is the occupation to society at large? How trendy is this occupation?

⁹ The median time required by participants was only 19 minutes. However, the standard deviation of 15 minutes indicates that some required significantly more time.

¹⁰ „Open AI’s text-davinci-003 completion model” (p. 8).

¹¹ Christoph^[75] carried out an update based on ISCO 1988. We consider Wegener’s version.

¹² The standard stimulus was the electrician, assigned a fixed value of 50.

¹³ The instruction was: „First, we would like to know your opinion on the PRESTIGE that various occupations currently have in Germany. Please assign each occupation a value from 0 to 10. A value of 0 indicates that the occupation has ‘very low prestige,’ and 10 indicates ‘very high prestige.’ You can use the values in between to nuance your opinion. In your opinion, how high is the prestige of <Occupation1-5> in Germany today?”

¹⁴ In random order, participants were successively presented with up to 80 facial photos of individuals of the opposite sex, each combined with random information about character, occupation, and body height. The task was to select the best partner for a lasting, long-term relationship. Preliminary choices could be replaced by new ones, but changes were limited to a maximum of four. Rejections could not be undone, and participants did not know how many potential partners would be offered.

¹⁵ In this case, where the same construct is being evaluated, the correlation itself corresponds to the true variance and therefore must not be squared.

¹⁶ Since the East German sample comprises 7 men and 6 women, there is no bias due to gender differences.

¹⁷ For the transformation of scales with different numbers of categories, various methods exist^[76]. For our correlational analyses, these differences are negligible.

¹⁸ The writer was not considered, as Bolte includes the specification „author of 50-pfennig novels” for this occupation.

¹⁹ The threshold of 0.75 standard deviation units is arbitrary. The value was chosen to ensure, on the one hand, a notable number of occupations and, on the other hand, results that remain manageable.

²⁰ The English-language publication by Huszár et al.^[27] refers only to occupational groups. The data on individual occupational titles are included as a supplement to the Hungarian-language article by Sik et al.^[28]. From this, we consider only the prestige values.

²¹ All other coefficients in the table are significant, and with two exceptions, $p < .001$.

²² Słomczyński and Kacprowicz^[77] and Domański, Sawiński and Słomczyński^[78] collected expert ratings for 500 and 551 occupations, respectively. These show, as expected with expert ratings, a much higher agreement with international scales. However, expert ratings are not relevant to our topic.

²³ For the politician, we used a value of 50 as a compromise between the parliament member (45) and government minister (54).

²⁴ For the politician, we used a value of 31.50 as a compromise between the parliament member (32.55) and the minister (30.38).

²⁵ The number 36 (three dozen) serves merely as an example; with 30 or 40, very similar results would be obtained, and even 20 raters would „guarantee” a highly reliable group standard.

²⁶ In magnitude scaling, which played a leading role in Germany for a long time but otherwise failed to gain traction, a different regularity must be considered. Here, variances in the lowest range are compressed, while those in the uppermost range are widely dispersed. For example, starting from the standard value of 50, half, a third, a quarter, and a fifth yield values of 25, 17, 13, and 10, respectively; while double, triple, quadruple, and quintuple yield 100, 150, 200, and 250. Thus, at the lower end, small absolute differences can be statistically significant, whereas at the upper end, large differences may be insignificant.

²⁷ According to our own calculations, the rainbow-shaped quadratic correlation between the rank position and the range of the 45 occupations from Counts is .83.

²⁸ The transformation is performed using the formula $f(x) = 100 - (\text{Maximum} - x) / (\text{Maximum} - \text{Minimum}) * 100$. It goes without saying that direct comparability is not possible, as the occupation lists are very different, and thus each occupation is evaluated in different contexts. Furthermore, it should be reiterated that there are various transformation methods for comparing scales with different numbers of categories^[76]. For our research question, these aspects are of little consequence.

²⁹ Both in 1983 in Beijing^[52] and in 1988 in Tianjin^[22], the greatest dispersion is found among high-ranking politicians.

³⁰ Ulfsdotter Eriksson^[54].

³¹ Already in the 1964/1989 comparison, the prestige of the firefighter had significantly increased^[14].

³² These figures are based on the 207 occupational groups from Garcia-Mainar and Montuenga^[23]. In the CIS^[24] with 285 occupations, the firefighter has the second-highest value.

³³ Here, a total of 21,515 participants in 28 countries evaluated the trustworthiness of 18 occupations.

³⁴ It should be noted, as elsewhere, that the correlations are based on different occupational samples.

³⁵ In comparing an early NORC study with prestige ratings from 23 countries, including many less developed ones, Poland showed the lowest agreement^[12], and in a comparison of 60 countries with the SIOPS, only 7 performed worse than Poland^[11].

³⁶ Ebner and Rohrbach-Schmidt^[38] report a correlation of .62 with the MPS and .72 with the SIOPS, describing these as „relatively strong,” although such low values are rarely found.

³⁷ Subjective judgments are not made in an absolute space but relative to a reference system. In our research focus on facial assessment, we have experimentally demonstrated striking context effects^[79]. Analogous effects were also shown in the assessment of the severity of offenses. It would be an intriguing task to replicate these experiments in the context of occupational prestige following the described approach.

³⁸ Doubts about the BAS also arise from the following finding: The test-retest reliability after 3 to 4 months is .93^[38]. The correlation between our scales is .974, despite a 9-year gap and the first survey being conducted face-to-face and the second as an online survey. By chance, exactly the same test-retest reliability was found in Sweden, as mentioned in the introduction, with a 16-year interval. However, it should be noted that the BAS covers more than four times as many occupations as our list and the Swedish list. Moreover, a test-retest reliability of .93 is more than sufficient.

³⁹ To establish a common denominator, one could also compile 5 sets differing in 75 or 76 occupations and sharing 25 occupations. This would increase the number of participants by a quarter.

⁴⁰ The gap between Whites and Blacks in the USA during the 1960s is undisputed, and despite enormous efforts, it has only slightly diminished to this day. In the 1960s, the educational level in the francophone part of Canada was significantly lower than in the anglophone part. A gradual convergence began only later. Slight differences in favor of anglophones are still evident in the most recent PISA study^[80]. There are significant overlaps between white-collar and blue-collar occupations, but it is beyond doubt that the average intelligence and educational level is higher in the white-collar group^{[81][82]}. The distinction between university, high school, and elementary education refers, by definition, to education.

⁴¹ „the teachers were somewhat more consistent in their judgments than were the high-school Seniors, since the average quartile range for the former is 6.0 and for the latter 8.8”^[7].

⁴² „disagreement in the relative ratings accorded to such functionaries as priests, high government officials ... may be the major consequence of cultural difference from country to country“^[12]. „we observe, as we move from country to country, that the relative standings of the most prestigious occupations such as religious officials ... and party or high government officials may change. These shifts have little impact upon the overall similarity in the prestige ratings of any pair of countries, but they almost surely harbor very important differences between countries as to sources of authority and control“ (^[83], S. 6).

⁴³ On our 0–100 scale, the standard deviation ranges from 8.5 to 28.6, with a median of 16.1 (Table 4). Overall, the disagreement in South Korea is much greater than in our study, but for a few occupations, opinions in our survey diverge more widely.

⁴⁴ Phenotypic intelligence is what an intelligence test measures. It correlates very highly with genetic potential but is not identical to it.

⁴⁵ This statement does not imply that humans tens of thousands of years ago had the same genetic potential as we do today. There is no doubt that the biological foundation has been improved in recent times through social selection.

⁴⁶ Contrary to what many believe, psychometric intelligence research is not biologicistic. On the contrary, it has always assumed that intelligence develops through the interaction of biology and culture, and a core concern has been to determine the weight of these two factors and their interaction. The Flynn Effect, uncovered by psychometric research, is a prime example of the significant role of social factors. Numerous critics of intelligence research fundamentally deny the role of biology, advocating a culturally deterministic stance that is justified by nothing other than ideology.

⁴⁷ In our opinion, the best book is „*The Development of Humanity from Childhood to Adulthood*“ by Georg W. Oosterdiekhoff^[84]. Additionally, we recommend Oosterdiekhoff^[85] and Hallpike^{[86][87]}. These works examine the interaction between social structure and intelligence from the perspective of Jean Piaget’s developmental psychology, which is indispensable for understanding the development of humanity.

⁴⁸ Lurija^[88] demonstrated in the 1930s through interviews in Uzbekistan, with many examples, that illiterate individuals are unable to solve even the simplest syllogisms, such as „*There are no camels in Germany. City B is in Germany. Are there camels there?*“ They cannot view the statements as a whole and instead point out that they have never been to Germany or City B. When faced with hypothetical questions that deviate from their own experiential world, they refuse to accept the premises.

References

1. ^a ^b ^c ^d ^e ^f ^g ^h ⁱ ^j ^k ^l Treiman DJ (1977). *Occupational Prestige in Comparative Perspective*. New York: Academic Press.
2. ^Δ Duncan OD (1961). "A Socio-Economic Index for All Occupations and Properties and Characteristics of the Socioeconomic Index." In Reiss AJ Jr (Ed.). *Occupations and Social Status*. Glencoe, IL: Free Press. S. 109–161.
3. ^Δ IfD (1965). *Am meisten geachtet: der Arzt. Schuldirektoren gewinnen an Ansehen. Eine neue Umfrage über die Einschätzung der Berufe*. Allensbach: Institut für Demoskopie.
4. ^Δ IfD (2024). *Berufsansehen*. Allensbacher Archiv, IfD-Umfrage 12082, Februar 2024. Allensbach: Institut für Demoskopie.
5. ^a ^b GfK Verein (2018). "Trust in Professions 2018 – A GfK Verein Study. From Firefighters to Politicians." Nürnberg: GfK Verein. Gesellschaft für Konsum-, Markt- und Absatzforschung e.V.
6. ^Δ YouGov (2021). "Scientists and Doctors Are the Most Respected Professions Worldwide." YouGov. <https://yougov.co.uk/international/articles/33901-scientists-and-doctors-are-most-respected-professions> [https://d3nkl3psvxxpe9.cloudfront.net/documents/YouGov - Profession_perceptions_international.pdf](https://d3nkl3psvxxpe9.cloudfront.net/documents/YouGov_-_Profession_perceptions_international.pdf).
7. ^a ^b ^c ^d ^e Counts GS (1925). "Social Status of Occupations." *Sch Rev*. 33:16–27.
8. ^a ^b ^c ^d ^e Smith M (1943). "An Empirical Scale of Prestige Status of Occupations." *Am Sociol Rev*. 8(2):185–192. doi:[10.2307/2085884](https://doi.org/10.2307/2085884).
9. ^Δ North CC, Hatt PK (1947). "Jobs and Occupations: A Popular Evaluation." *Opin News*. 3–13.
10. ^a ^b ^c Hodge RW, Siegel PM, & Rossi PH. (1966). Occupational prestige in the United States, 1925–63. In: Bendix R, & Lipset SM. (Eds.). *Class, status, and power. Social stratification in comparative perspective*. Second Edition. New York: The Free Press, p. 322–334.
11. ^Δ Inkeles A, Rossi PH (1956). "National Comparisons of Occupational Prestige." *Am J Sociol*. 61(4):329–339. <https://www.journals.uchicago.edu/doi/abs/10.1086/221764>.
12. ^a ^b ^c Hodge RW, Treiman DJ, Rossi PH (1966). "A Comparative Study of Occupational Prestige." In Bendix R, Lipset SM (Eds.). *Class, Status and Power*. New York: Free Press, p. 309–321.
13. ^a ^b ^c Nakao K, Treas J (1990). "Computing 1989 Occupational Prestige Scores." *GSS Methodological Report No. 70*.
14. ^a ^b ^c ^d ^e ^f Nakao K, Treas J (1994). "Updating Occupational Prestige and Socioeconomic Scores: How the New Measures Measure Up." *Sociol Methodol*. 24:1–72.

15. ^ΔNakao, K. (n.d.). Occupational Prestige. *Encyclopedia of Sociology*. Retrieved August 15, 2024 from Encyclopedia.com: <https://www.encyclopedia.com/social-sciences/encyclopedias-almanacs-transcripts-and-maps/occupational-prestige>
16. ^ΔNakao, K. & Treas, J. (1992). The 1989 Socioeconomic Index of Occupations: Construction from the 1989 prestige scores. *GSS Methodological Report No. 74*.
17. ^ΔSiegel, P.M. (1971). *Prestige in the American occupational structure*. Unpublished Ph.D. dissertation, University of Chicago.
18. ^ΔSmith TW, Son J (2014). "Measuring Occupational Prestige on the 2012 General Social Survey." *GSS Methodological Report, No. 122*. Chicago: NORC.
19. ^ΔPohoski M, Słomczyński KM, Wesołowski W (1976). "Occupational Prestige in Poland, 1958-1975." *Pol Sociol Bull.* 36:63–77. <http://www.jstor.org/stable/44824694>.
20. ^{a, b, c, d, e, f, g, h}Domański H, Sawiński Z, Słomczyński KM (2010). "Prestiż zawodów w obliczu zmian społecznych: 1958-2008" [Occupational Prestige Under Social Change: 1958-2008]. *Stud Socjol.* 4(199):79–119.
21. ^{a, b}Chiu CCH (1999). "The Changing Occupations and Their Prestige Structure in Urban China." *China Rev.* 30 9–328. <http://www.jstor.org/stable/23453325>.
22. ^{a, b}Bian Y (1996). "Chinese Occupational Prestige. A Comparative Analysis." *Int Sociol.* 11(2):161–186. doi:[10.1177/026858096011002002](https://doi.org/10.1177/026858096011002002).
23. ^{a, b}Garcia-Mainar I, Montuenga VM (2025). "Is the „Treiman Constant” Actually Constant? An Assessment Using Two Spanish Occupational Prestige Scales: 1991 and 2013." *Soc Sci Res.* 126:103125. doi:[10.1016/j.ssresearch.2024.103125](https://doi.org/10.1016/j.ssresearch.2024.103125).
24. ^{a, b, c, d, e, f, g}CIS (2013). "Prestigio ocupacional y estructura social" [Occupational Prestige and Social Structure]. Estudio 3004. Centro de Investigaciones Sociológicas (CIS).
25. ^{a, b, c}Ulfsdotter Eriksson Y, Nordlander E (2024). "Occupational Prestige and Gendered Polarisation." In Berglund T, Ulfsdotter Eriksson Y (Eds.). *Scrutinising Polarisation: Patterns and Consequences of Occupational Transformation in the Swedish Labour Market*. London: Routledge. S. 56-72. doi:[10.4324/9781003412861](https://doi.org/10.4324/9781003412861).
26. ^{a, b}Hughes BT, Srivastava S, Leszko M, Condon DM (2024). "Occupational Prestige: The Status Component of Socioeconomic Status." *Collabra Psychol.* 10(1). doi:[10.1525/collabra.92882](https://doi.org/10.1525/collabra.92882).
27. ^{a, b}Huszár Á, Hajdu G, Sik E, Nagy RK (2023). "New Hungarian Occupational Prestige Scale." *Hung Stat Rev.* 6 (2):12–36. doi:[10.35618/HSR2023.02.en012](https://doi.org/10.35618/HSR2023.02.en012).
28. ^{a, b, c}Sik E, Huszár Á, Nagy KR (2023). "A Foglalkozások Inkonzisztenciája – Egy Új, Presztízsalapú Tipológia Felé" [The Inconsistency of Occupations – Towards a New, Prestige-Based Typology]. *Statiszt Szemle.* 101(2):

- ge of Occupations – Construction of a New Occupational Prestige Scale and Empirical Findings for Germany]. *Z Soziol.* 50(6):349–372. doi:[10.1515/zfsoz-2021-0026](https://doi.org/10.1515/zfsoz-2021-0026).
39. [△]Stevens, S.S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects*. New York: Wiley.
 40. [△]Henss R (1987). "Zur Beurteilerübereinstimmung bei der Einschätzung der Physischen Attraktivität Junger und Alter Menschen" [On Rater Agreement in the Assessment of the Physical Attractiveness of Young and Old People]. *Z Sozialpsychol.* 18:118–130.
 41. [△]Henss, R. (1991). Perceiving age and attractiveness in facial photographs. *Journal of Applied Social Psychology*, 21,933–946.
 42. [△]^bHenss R (1992). „Spieglein, Spieglein an der Wand ...“ Geschlecht, Alter und Physische Attraktivität [“Mirror, Mirror on the Wall ...” Gender, Age and Physical Attractiveness]. Weinheim: Psychologie Verlags Union.
 43. [△]^bHenss R (1998a). *Gesicht und Persönlichkeitseindruck* [Face and Personality Impression]. Göttingen: Hogrefe. Reprint: Henss R (2020). *Gesicht und Persönlichkeitseindruck*. Saarbrücken: Dr. Ronald Henss Verlag. https://www.researchgate.net/publication/339147401Gesicht_und_Persönlichkeitseindruck.
 44. [△]Henss R (1998b). "Type Nouns and the Five Factor Model of Personality Description." *Eur J Pers.* 12:57–71.
 45. [△]Henss, R. (1997). Face and personality impressions. Consensus at zero acquaintance. Poster presented at the 4. Arbeitstagung Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik. University of Bamberg, Germany. Oct. 30 – Nov. 01, 1997. <https://doi.org/10.13140/RG.2.2.11895.23205>
 46. [△]Breinig RM (1999). *Multiple Kriterien der Partnerwahl. Was Männer wünschen* [Multiple Criteria for Mate Choice. What Men Want]. Unveröffentlichte Diplomarbeit. Fachrichtung Psychologie. Universität des Saarlandes, Saarbrücken.
 47. [△]Schmidt IM (2001). *Multiple Kriterien der Partnerwahl. Was Frauen Wünschen* [Multiple Criteria for Mate Choice. What Women Want]. Unveröffentlichte Diplomarbeit. Fachrichtung Psychologie. Universität des Saarlandes, Saarbrücken.
 48. [△]Goldthorpe, J.H. & Hope, K. (1974). *The social grading of occupations. A new approach and scale*. Oxford: Clarendon Press.
 49. [△]Davies AF (1952). "Prestige of Occupations." *Br J Sociol.* 3(2):134–147.
 50. [△]Hout M, DiPrete TA (2006). "What We Have Learned: RC28's Contributions to Knowledge About Social Stratification." *Res Soc Stratif Mobil.* 24(1):1–20. doi:[10.1016/j.rssm.2005.10.001](https://doi.org/10.1016/j.rssm.2005.10.001).
 51. [△]Buss DM (2016). *The Evolution of Desire: Strategies of Human Mating*. New York: Basic Books.

52. ^aLin N, Xie W (1988). "Occupational Prestige in Urban China." *Am J Sociol.* 93(4):793–832. <http://www.jstor.org/stable/2780604>.
53. ^aUlfssdotter Eriksson Y (2024). "Polarised Perceptions of Occupational Prestige?" In Berglund T, Ulfssdotter Eriksson Y (Eds.). *Scrutinising Polarisation: Patterns and Consequences of Occupational Transformation in the Swedish Labour Market*. London: Routledge. S. 73–86. doi:[10.4324/9781003412861](https://doi.org/10.4324/9781003412861).
54. ^aUlfssdotter Eriksson, Y. (2006). *Yrke, Status & Genus – en sociologisk studie om yrken på en segregerad arbetsmarknad*. [Occupation, Status & Gender – A sociological study of occupations on a segregated labour market]. Doctoral Dissertation at the Department of Sociology, Göteborg University.
55. ^ΔValentino L (2022). "Status Lenses: Mapping Hierarchy and Consensus in Status Beliefs." *RSF J Soc Sci.* 8(7): 89–110. <https://www.jstor.org/stable/48703117>.
56. ^ΔJacobs J, Powell B (1984). "Gender Differences in the Evaluation of Prestige." *Sociol Q.* 25(2):173–190.
57. ^ΔKrueger S, Ebner C, Rohrbach-Schmidt D (2022). "Gender Composition and the Symbolic Value of Occupations: New Evidence of a U-Shaped Relationship Between Gender and Occupational Prestige Based on German Microdata." *Work Employ Soc.* 38(1):242–261. doi:[10.1177/09500170221117415](https://doi.org/10.1177/09500170221117415).
58. ^ΔWurzbacher, G. (1954). *Das Dorf im Spannungsfeld industrieller Entwicklung*. Stuttgart: Ferdinand Enke Verlag.
59. ^ΔIpsos (2022). "Global Trustworthiness Index 2022. Who Does the World Trust?" Ipsos. <https://www.ipsos.com/sites/default/files/ct/news/documents/2022-07/Global%20trustworthiness%202022%20Report.pdf>.
60. ^ΔBerglund, T. & Ulfssdotter Eriksson, Y. (Eds.) (2024). *Scrutinising polarisation. Patterns and consequences of occupational transformation in the Swedish labour market*. Abingdon, Oxon: Routledge. Routledge Studies in the Sociology of Work, Professions and Organisations, <https://doi.org/10.4324/9781003412861>
61. ^ΔGuppy LN (1984). "Dissensus or Consensus: A Cross-National Comparison of Occupational Prestige Scales." *Can J Sociol / Cah Can Sociol.* 9(1):69–83. doi:[10.2307/3340468](https://doi.org/10.2307/3340468).
62. ^ΔPineo, P.C. & Porter, J. (1967). Occupational prestige in Canada. *Canadian Review of Sociology and Anthropology*, 4, 24–40.
63. ^ΔHaier RJ, Colom R, Hunt E (2024). *The Science of Human Intelligence. Second Edition*. Cambridge: Cambridge University Press.
64. ^aHenss R (2025b). *The Intelligence of Nations. National IQs and Correlates*. Saarbrücken: Dr. Ronald Henss Verlag. Also at Qeios, 2024. doi:[10.32388/6IBSKW](https://doi.org/10.32388/6IBSKW).
65. ^ΔRindermann H (2018). *Cognitive Capitalism. Human Capital and the Wellbeing of Nations*. Cambridge: Cambridge University Press.

66. [△]Weiss V (2012). *Die Intelligenz und Ihre Feinde. Aufstieg und Niedergang der Industriegesellschaft [Intelligence and Its Enemies. Rise and Fall of Industrial Society]*. Graz: Ares Verlag.
67. [△]Warne, R.T. (2020). *In the know. Debunking 35 myths about human intelligence*. Cambridge: Cambridge University Press.
68. [△]Flynn JR (2009). *What Is Intelligence? Beyond the Flynn Effect*. Cambridge: Cambridge University Press.
69. [△]Flynn JR (2012). *Are We Getting Smarter? Rising IQ in the Twenty-First-Century*. Cambridge: Cambridge University Press.
70. [△]Henss R (2025c). *Die Global Learning Crisis. Schulbesuch ist Nicht Gleich Lernen [The Global Learning Crisis. School Attendance Is Not the Same as Learning]*. Saarbrücken: Dr. Ronald Henss Verlag. Also at ResearchGate, 2023: <https://doi.org/10.13140/RG.2.2.11286.83520>.
71. [△]Kohlberg L (1996). *Die Psychologie der Moralentwicklung [The Psychology of Moral Development]*. Frankfurt/Main: Suhrkamp.
72. [△]Piaget J (1973). *Das Moralische Urteil beim Kinde [The Moral Judgment of the Child]*. Frankfurt/Main: Suhrkamp.
73. [△]Henss R (2025a). "Ein Etwas Anderer Blick auf Berufsprestigeratings" [A Slightly Different Look at Occupational Prestige Ratings]. Saarbrücken: Dr. Ronald Henss Verlag. Also at ResearchGate, 2025: <https://doi.org/10.13140/RG.2.2.21699.31528>.
74. [△]Wikipedia (n.d). *Liste der häufigsten Familiennamen in Deutschland*. https://de.wikipedia.org/wiki/Liste_der_h%C3%A4ufigsten_Familiennamen_in_Deutschland
75. [△]Christoph B (2005). "Zur Messung des Berufsprestiges: Aktualisierung der Magnitude-Prestigeskala auf die Berufsklassifikation ISCO88" [On the Measurement of Occupational Prestige: Updating the Magnitude Prestige Scale to the ISCO88 Occupational Classification]. *ZUMA Nachr.* 29(57):79–127. <https://nbn-resolving.org/urn:nbn:de:0168-ssnar-207543>.
76. [△]^bHenss R (1989). "Zur Vergleichbarkeit von Ratingskalen mit Unterschiedlicher Kategorienzahl" [On the Comparability of Rating Scales with Different Numbers of Categories]. *Psychol Beitr.* 31(3/4):264–284.
77. [△]Ślomoński KM, Kacprowicz G (1979). *Skale Zawodów [Occupational Scales]*. Warszawa: Polska Akademia Nauk, Instytut Filozofii i Socjologii.
78. [△]Domański H, Sawiński Z, Ślomoński KM (2009). *Sociological Tools Measuring Occupations: New Classification and Scales*. Warsaw: IFiS Publishers.
79. [△]Henss R (1993). "Kontexteffekte bei der Beurteilung der Physischen Attraktivität" [Context Effects in the Assessment of Physical Attractiveness]. In Hassebrauck M, Niketta R (Hrsg.). *Physische Attraktivität*. Göttingen: H

ogrefe. S. 61-94.

80. [△]CMEC (2023). "Measuring Up: Canadian Results of the OECD PISA 2022 Study. The Performance of Canadian 15-Year-Olds in Mathematics, Reading, and Science." Toronto: Council of Ministers of Education, Canada. <https://www.cmec.ca/Publications/Lists/Publications/Attachments/438/PISA-2022 Canadian Report EN.pdf>.
81. [△]Gottfredson LS (1986). "Occupational Aptitude Patterns Map: Development and Implications for a Theory of Job Aptitude Requirements." *J Vocat Behav.* 29(2):254–291. doi:10.1016/0001-8791(86)90008-4.
82. [△]Gottfredson LS (1997). "Why g Matters. The Complexity of Everyday Life." *Intelligence.* 24:79–132.
83. [△]Nakao, K., Hodge, R.W. & Treas, J. (1990). On revising prestige scores for all occupations. GSS Methodological Report No. 69. Chicago: The University of Chicago.
84. [△]Oesterdieckhoff GW (2013). *Die Entwicklung der Menschheit von der Kindheitsphase zur Erwachsenenreife [The Development of Humanity from Childhood to Adulthood]*. Wiesbaden: Springer VS.
85. [△]Oesterdieckhoff, G.W. (2012). *Die geistige Entwicklung der Menschheit*. Weilerswist: Velbrück.
86. [△]Hallpike, C.R. (1979). *The foundations of primitive thought*. Oxford: Clarendon Press.
87. [△]Hallpike, C.R. (2008). *How we got here. From bows and arrows to the space age*. Milton Keynes: AuthorHouse.
88. [△]Luria, A.R. (1976). *Cognitive development: Its cultural and social foundations*. Cambridge, MA: Harvard University Press.

Supplementary data: available at <https://doi.org/10.32388/JG9O38>

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.