Exploring occupational prestige through Large Language Models: A multi-dimensional approach

Robert de Vries[1] (corresponding author)

Mark Hill[2]

Laura Ruis[3]

[1] University of Kent, School of Social Sciences

[2] Kings College London, Department of Digital Humanities

[3] Massachusetts Institute of Technology, Language and Intelligence Lab

Keywords: stratification, prestige, social status, occupation, large language models, ChatGPT

Declaration of interest: none

Data statement: the data and code on which our reported analysis is based are available from https://doi.org/10.7910/DVN/2HI67X

Funding statement: This project was partly funded by the University of Kent Division of Law, Society, and Social Justice Investment in Research Fund.

Formal published text DOI: https://doi.org/10.1016/j.rssm.2026.101185

Abstract

Debates over occupational prestige have long centred on what existing prestige scales actually measure and whether prestige constitutes a single evaluative hierarchy or multiple distinct dimensions. This paper contributes to these debates both conceptually and methodologically by examining the dimensionality of occupational prestige as represented in the semantic associations learned by Large Language Models (LLMs) from human text. Using pairwise comparisons across a representative list of occupations, we prompt multiple LLMs to generate rankings along five scales: status, prestige, social standing, expected deference, and an entirely novel dimension based on advertised social proximity (as an indicator of status-related associational preference bias).

We show that status, prestige, and social standing are empirically indistinguishable in the models' associative structure, forming a unified hierarchy closely aligned with established survey-based prestige scales. By contrast, deference and advertised proximity emerge as clearly separable dimensions. Further analyses demonstrate that these dimensions are differentially related to associations the models hold about occupational characteristics such as training and pay, authority, and social contribution. Substantively, the findings support a multidimensional conception of occupational prestige as represented in LLMs' associative structure – and specifically suggest advertised social proximity as a novel, unexplored dimension. Methodologically, we demonstrate – while recognising their limitations – the value of LLM-based approaches as a complement to human research for exploring complex evaluative structures that are difficult to measure using conventional survey methods.

1 Introduction

Which occupations are looked up to and which are looked down on is an essential question in the study of social stratification. Social status is a key dimension of social hierarchy (Chan & Goldthorpe, 2007), and as Chan (2010) argues "in modern societies occupation is one of the most salient positional characteristics to which status attaches" (p.29). A person's occupation is often one of the first personal details we learn about them, helping us to socially 'place' them relative to ourselves and others (Weeks & Leavitt, 2017).

There are a variety of existing measures of occupational prestige, including the Standard International Occupational Prestige Scale (SIOPS) (Treiman, 1977), the 2012 US General Social Survey (GSS) scale (Smith & Son, 2014), and the recently introduced Newlands and Lutz (2024) scale. However, there remains widespread disagreement about exactly what occupational prestige scales capture, or should aim to capture. For example, there is longstanding debate around whether existing survey measures tap prestige in the Weberian sense of "a social assessment of honour" (Weber, 2010), or only an occupation's general 'desirability' (based principally on training requirements and financial rewards) (Bukodi et al., 2011). Do lawyers perform well on survey-based prestige scales because they are genuinely widely esteemed and respected, or simply because respondents know that 'lawyer' is considered a 'good' job because it is well-paid and requires extensive training? This question is particularly relevant where surveys employ general terms like 'social standing' (Goldthorpe & Hope, 1972). However, even where the specific language of 'prestige' is adopted, it is not entirely clear that respondents will interpret this in the strict Weberian sense (Valentino, 2021).

This issue is exacerbated by the fact that respondents' prestige judgements are, to an extent, 'second order' attitudes. Responses to whether, for example, 'stockbroker' is an occupation that has "high social standing" are likely to result from a complex interplay between the respondent's own personal views, and their sense of the attitudes of wider society. I may base my personal view of the 'social standing' of occupations principally on their contribution to society, and consequently score stockbrokers relatively poorly (Ulfsdotter Eriksson & Nordlander, 2023). However, I may also be aware that not everyone shares my view of what confers standing, and adjust my response accordingly. The extent of this adjustment – and therefore exactly what attitude is being captured – is likely to vary from person to person (Valentino, 2021).

Some authors have argued that traditional survey approaches therefore do not adequately capture occupational prestige as a concept (Bukodi et al., 2011). For example Freeland and Hoey (2018) make the case for assessing prestige on the basis of deference relationships. They argue that social status is fundamentally rooted in (citing Podolny and Lynn, 2013) "accumulated acts of deference". On this view, the occupational prestige order is essentially the order of occupations according to who would generally defer to whom (Freeland & Hoey, 2018).

One could argue, however, that the act of deference, rather than solely indicating relationships of respect and esteem, is instead strongly influenced by formal arrangements of power and authority. For example, people at all levels in the hierarchy of respect are legally required (or normatively expected) to defer to police officers in many contexts. Notably, Zhou (2005) argues that such authority relationships are an inherently fragile basis for claims to prestige, since they are always open to contest and challenge – unlike claims rooted in 'nature and reason'.

A similar argument to that of Freeland and Hoey (2018) could be developed to support the measurement of prestige through patterns of *associational preference bias* rather than deference. It is well established that people prefer to socially associate with, and advertise their social proximity to, high status others (Thye, 2000; Sauder, et al., 2012) – primarily because, as Ridgeway (2014) explains, "the status of those with whom an actor associates affects that actor's own status…(i.e. status 'spreads' through association)" (p.6). Indeed, one could say that the easiest way to identify the highest status person in a room is to see who has attracted the biggest crowd of attempted interlocutors; or the largest number of claims to close association. Capturing the associational preference bias associated with specific occupations could therefore provide profound insights into the occupational prestige order. To our knowledge, this is not a possibility that has been explored in the literature thus far.[1]

At root, these debates concern the construct validity and dimensional structure of occupational prestige. What, precisely, do existing occupational prestige scales capture? Does there exist a single, underlying prestige order, on which these measures offer variously distorted views? Or is it more helpful to conceptualise prestige in terms of multiple, separate evaluative dimensions – each producing equally valid hierarchies of occupations?

A sensible approach to answering these questions would be to examine patterns of correlation between multiple measures of prestige and potentially related evaluative dimensions. How, for example, are deference relationships and associational preference bias related to each other, and to perceptions of social standing? In what ways do measures of these concepts diverge, and what do these divergences reveal?

Previous studies have attempted some comparisons of this nature. For example, Freeland and Hoey (2018) compare their deference scale to existing survey-based scales employing the 'social standing' and 'prestige' language – finding that deference is more strongly related to the latter than the former. Similarly, Newlands and Lutz (2024) compare their occupational prestige measure with a measure of perceived social value – finding that these measures are highly correlated, but with illuminating patterns of divergence. The scope of this type of comparative work is, however, strongly constrained by practical considerations. Human participants do not have infinite attention spans, and cannot be expected to rank or score hundreds of occupations on large numbers of scales. This is particularly the case for measures of deference, which inherently involve pairwise contrasts – an issue we return to below. Survey burden can be reduced by using sub-samples (Smith & Son, 2014; Newlands & Lutz, 2024). However, this creates issues of statistical precision, and militates against many kinds of analyses that would require multiple measures from the same respondents. These issues are exacerbated by the potential for even small changes in terminology (for example, the difference between 'prestige', 'respect', 'social status', and 'social standing') to have an outsize effect on responses. Hence existing research has been limited to comparisons between a small number of scales (e.g. Newlands & Lutz, 2024), or with external scales produced using vastly different approaches and populations (e.g. Freeland & Hoey, 2018).

[1] Note that the logic here is different from that underlying occupational status/prestige measures based on patterns of social association, such as the Chan-Goldthorpe (2004) scale. These scales are based on the expectation that close social relationships (such as marriage and close friendship) form most easily among status peers. The formation of these *mutual* close bonds involves different mechanisms than a preference to associate oneself with high status others.

In this paper, we argue that Large Language Models (LLMs) can make a contribution in addressing these questions. We show that, through careful prompting, the semantic associations embedded within LLMs produce multiple empirically and conceptually separable evaluative dimensions related to occupational prestige. We also show how this approach can be used to explore how important occupational characteristics – such as perceived training requirements and pay – may differentially predict performance on these dimensions.

1.1 Using Large Language Models to measure occupational prestige

LLMs such as ChatGPT (OpenAI, 2023), Claude (Anthropic, 2024), and Llama (Touvron et al., 2023) are probabilistic models of text created through a 'training' process, one important objective of which is to correctly predict the next word in a piece of text. Modern LLMs are trained on truly enormous databases, comprising text scraped from large fractions of the web; as well as from academic articles, digitised books, and computer code (Gao et al., 2020). Through the process of training on large amounts of text, representations of associations between textual elements emerge. For example, a model will learn that buses are very commonly referred to as 'late' and very rarely referred to as 'tasty'. Research has shown that in modern LLMs, these associations are based on *semantic* relationships between words rather than their simple co-occurrence in language (Digutsch & Kosinski, 2023). For example, LLMs will associate 'mountain' more strongly with 'high' than with 'top', despite the latter being more common as the next word in a sentence (Digutsch & Kosinski, 2023).

Because the associations LLMs hold – captured in word embeddings – are based on text written (predominantly) by humans, they tend to reflect societal biases and stereotypes (Caliskan et al., 2017). For example, in human-written text about medical professionals, doctors are most often men, while nurses tend to be women (Kirk et al., 2021). Hence, absent explicit tuning to the contrary (Raza et al., 2024), LLMs will hold a strong semantic association between the concepts 'doctor' and 'man/male'. These biases are detectable when prompting the model. For example, research has shown that LLMs tend to assume that lawyers and doctors are men (Kirk et al., 2021), that they more strongly associate violence related concepts with African American than with European American women (An et al., 2023), and that they strongly associate terrorism and violence related concepts with Muslims (Abid et al., 2021).

Biases and stereotypes in LLMs are most often researched in the context of their potentially problematic social consequences – for example in perpetuating gender and racial stereotypes (Kirk et al., 2021; An et al., 2023). However, in that they derive from statistical associations in human-written training data, they may also offer a window into *human* biases and stereotypes. For example, if LLMs tend to assume that doctors and lawyers are men, it is because this bias exists in the text on which the model was trained. Given that this constitutes a large fraction of all internet text, this is a strong indicator of a real societal bias. The logic of using large language models in this way is well-captured by technologist Jeff Jarvis in his submission to a 2024 US Congressional hearing on AI: "Because LLMs are in essence a concordance of all available language online, I hope to see scholars examine them to study society's biases and clichés." (Jarvis, 2024). Here we aim to apply this notion specifically to examining the dimensionality of occupational prestige.

In their training, LLMs will have developed associations between occupational titles and prestige-related words and concepts. For example, in books, news articles, and social media posts, lawyers and doctors are more likely to be described in prestigious terms than are opticians or office managers. Similarly, one would expect police officers to be more commonly

described as receiving social deference than postal workers or shop assistants. These statistical association will be encoded in weights in the models' 'embedding space' (Digutsch & Kosinski, 2023). In this study, we attempt to recover the strength of these associations through prompting, and thereby examine how prestige and related dimensions are represented in the model's associational structure.

The possibility that LLMs may offer a useful insight into human attitudes has been explored in a number of recent studies (Argyle et al., 2022; Sarstedt et al., 2024; Kim & Lee, 2024). However, there are substantial concerns with using LLMs in this way. These concerns can be attributed to the two inferential steps required to attribute social meaning to LLM outputs.

The first concerns whether LLM outputs reliably capture the actual associations present in the model's embedding space – and by extension, in the human-written text corpus on which the model was trained. Most LLMs do not respond directly from "raw" word embeddings or from pretraining alone. Their outputs reflect multiple stages of development, including instruction tuning, human-feedback-based alignment, and safety filtering, which are often intended to reduce socially or commercially problematic responses (Santurkar et al., 2023). As a result, a prompted response may reflect, in addition to training-data associations; prompt effects, alignment interventions, or deployment-specific filtering. Studies using LLMs to simulate survey responses show that chat model outputs can be brittle: the same prompt may yield different answers over time, while small wording changes can produce large response shifts (Bisbee et al., 2024). Outputs from different models may also vary substantially because of differences in training data, architecture, fine-tuning, alignment, or inference settings.

The second inferential step concerns whether the textual associations in training data – if they can indeed be reliably recovered through prompting – are informative about real human attitudes. Here a key concern is analogous to survey sampling bias. Most contemporary LLMs are trained on text that is dominated by English-language content produced by and for Western populations (Santurkar et al., 2023). They may therefore better reflect the perceptions of these than other populations (Santurkar et al., 2023).

A further source of bias is that online text represents only one slice of human experience. Attitudes expressed in text – even in informal contexts – may differ systematically from those expressed in other modalities (Kozlowski & Evans, 2025). This potential divergence is suggested by research showing that LLMs can exhibit, for example, sexist biases *more strongly* than humans (Gallegos et al., 2024) – potentially because online text contains more stereotypical depictions of women than are reflected in real everyday experience.

Any study which attempts to draw even tentative conclusions about human attitudes from LLM outputs must account for both of the above inferential steps. The next section explains how our methodology attempts to address the first (from model outputs to textual associations). The second more speculative step, from textual associations to human attitudes, is taken up at length in the Discussion.

1.2 The present study

In the following sections we report two sets of analyses. First, we attempt to explore the dimensional structure of occupational prestige, as captured in the semantic associations embedded in LLMs. We employ three large models: two commercial, closed-weights models (OpenAI's GPT-4o and GPT4.1); and one open-weights model (Meta's Llama 3.3 70B Instruct).

Do the associations embedded in these models appear to reflect a single latent prestige order, or a set of clearly separable evaluative dimensions?

We first examine how the model ranks occupations according to three terms commonly used in surveys of occupational prestige: 'prestige, 'status', and 'social standing'. We next analyse how the same occupations were ranked according to the two additional prestige-related dimensions discussed above: **deference** (the likelihood that one occupation would defer to another), and **associational preference bias**. As we have noted, associational preference bias is often reflected through 'advertised proximity', whereby people highlight their social proximity to high-status others in order to increase their own status by association (Ridgeway, 2013). We therefore attempted to capture this concept by asking the model to contrast occupations by how likely someone would be to boast about meeting a person in that occupation at a social event (see specific prompts below).

In our second set of analyses, we examine the extent to which these three prestige dimensions are differentially predicted by the following characteristics of occupations, again as rated by the model:

- **Training requirements and pay**. Previous research has argued that perceived training requirements and pay are key determinants of prestige judgements (Adler & Kraus, 1985). As noted above, others have argued that existing prestige scales are mis-named and in fact capture *only* perceptions of job quality as determined by requirements and rewards (Goldthorpe & Hope, 1972). Our analysis therefore explores whether the model's prestige rankings can be entirely explained by associations related to requirements and rewards.

- **Social contribution**. Existing scholarship has taken two competing positions on the role of social contribution in occupational prestige. Either it is a determinant of prestige (Mackinnon & Langford, 1994), or it is itself a part of the prestige concept (Freeland & Hoey, 2018). Here we take the former view and examine social contribution as a potential influence on prestige perceptions.

- **Authority relations**. Previous research has argued that authority relations are a secondary determinant of prestige judgements (Zhou, 2005). Here we have a clear expectation, as articulated above, that authority relations play a strong role in explaining patterns of *deference*, and are less influential on more general perceptions of social standing.

- **'Unusualness', public visibility, excitement and interest.** To our knowledge, previous research has not directly explored these factors as determinants of occupational prestige. However, broader research suggests that perceptions of how unusual, interesting/exciting, or (especially) publicly visible an occupation is shape judgements of job desirability (Hoff et al., 2024). This suggests that they may also play a role in occupational prestige judgements.

Three aspects of our approach make us confident that our results provide a valid insight into the models' underlying embedding space, and therefore into the text on which the models were trained. First, we replicate our analyses across three different models from two different companies. A robust pattern of results across all models would suggest that our findings are not an artefactual product of the interaction between a specific prompt and a specific model.

Second, our approach differs from previous 'silicon sampling' approaches, which typically ask models to adopt socio-demographic 'personas' and to answer a set of specific survey questions (e.g. Argyle et al., 2022; Bisbee et al., 2024). As noted above, such approaches can provide brittle, time and model dependent results (Bisbee et al., 2024). By contrast, we ask our models about general social perceptions – for example, about whether an occupation is "generally perceived" to be prestigious (see Section 2.3, below). This is more akin to asking, for example, for the meaning of a word as it is generally used. This likely gives better access to the model's underlying associational structure than asking it to give e.g. a political opinion while posing as a person with specific socio-demographic statistics.

Third, our conclusions are based on the overall *pattern* of associations we observe – particularly concerning the dimensionality of prestige. While responses to a single, specific prompt may be brittle, the pattern of associations between dimensions should be more robust, and more reflective of the underlying associations present in the training data.

Overall, we find that, within all three models, occupational rankings of prestige, status, and social standing are almost perfectly correlated – suggesting that these terms capture a unified latent concept in the models' associational network. By contrast, rankings of deference and advertised proximity were clearly separable, with distinct and intuitive clusters of discrepant occupations – for example, police officers (moderate status, high likelihood of deference) versus actors (higher status, low likelihood of deference); and social media influencers (low status, high advertised proximity) versus accountants (high status, low advertised proximity). Again, this pattern was consistent across all three models. We also found that these three dimensions (status/prestige/standing, deference, and advertised proximity) were associated in sensible ways with model 'perceptions' of occupational characteristics – for example, deference was more closely related to authority relations, while advertised proximity was more strongly associated with features of occupations that make them unusual, interesting, or exciting.

Taken together, or results suggest that LLMs trained on human-produced text encode a stable, interpretable, and externally plausible multidimensional structure of prestige evaluation. As we argue above, this strongly suggests that this associational structure is present in the underlying text on which the model was trained. In the Discussion, below, we discuss the implications of this finding, including the extent to which it can offer insights into human prestige judgements.

2 Methods

2.1 Occupations

We constructed a list of 99 occupations for comparison, chosen to broadly represent all 43 sub-major groups present in ISCO-08. We further included two occupations which share the same occupational code (*Waiter* and *Waitress*) to ascertain whether the model recognised prestige differences not captured by existing scales. Formal occupation titles were translated to colloquial US English terms (for example, 'paramedical practitioner' was translated to 'paramedic').

The full list of occupations, along with their corresponding ISCO-08, OCC 2000 (IPUMS, 2000a), and OCCSOC 2000 (IPUMS, 2000b) codes and labels are given in Appendix A.

2.2 Models

The prompts described below were passed to three models:

- OpenAI's gpt-4o and gpt-4.1 using the OpenAI API.
- Meta's llama_v3p3_70b_instruct through the Fireworks AI API.

No additional system-level prompts were specified and no tools were attached (the models therefore could not search the web, for example). Model temperature was set at 0 to encourage the model to consistently return only the most probable result. Model temperature refers to how the model chooses the subsequent token in an output (OpenAI, 2024). Lower values restrict the model to the highest probability tokens, with a value of 0 making its outputs largely deterministic (Ruis et al., 2023).

2.3 Prompting approach

The results reported below are based on the following five core prompts (/n indicates a line break):

Status

There are two occupations: /n 1) [OCCUPATION 1]./n 2) [OCCUPATION 2]. /n Which occupation is generally perceived to be higher status? Answer only using the numbers 1 ([OCCUPATION 1] is generally perceived to be higher status) or 2 ([OCCUPATION 2] is generally perceived to be higher status)

Prestige

There are two occupations: /n 1) [OCCUPATION 1]./n 2) [OCCUPATION 2]. /n Which occupation is generally perceived to be more prestigious? Answer only using the numbers 1 ([OCCUPATION 1] is generally perceived to be more prestigious) or 2 ([OCCUPATION 2] is generally perceived to be more prestigious])

Social standing

There are two occupations: /n 1) [OCCUPATION 1]./n 2) [OCCUPATION 2]. /n Which occupation is generally perceived to have higher social standing? Answer only using the numbers 1 ([OCCUPATION 1] is generally perceived to have higher social standing) or 2 ([OCCUPATION 2] is generally perceived to have higher social standing)

Deference

Please choose between the following two statements. In a general social situation, is it more likely that /n 1) a [OCCUPATION 1] would defer to a [OCCUPATION 2] or /n 2) a [OCCUPATION 2] would defer to a [OCCUPATION 1]? /n Answer only using the numbers 1 ([OCCUPATION 1] defers to [OCCUPATION 2]) or 2 ([OCCUPATION 2] defers to [OCCUPATION 1])

Advertised proximity

A social event has happened. Riley[2] has had the opportunity to chat to a variety of different kinds of people at this event, including /n 1) a [OCCUPATION 1], and /n 2) a [OCCUPATION 1]. /n Which person is Riley most likely to boast about making friends with? Answer only using the

[2] Riley is the most gender neutral first name in the US, with a male:female ratio of 51:49 (Flowers, 2015).

numbers 1 (Riley is more likely to boast about making friends with [OCCUPATION 1]) or 2 (Riley is more likely to boast about making friends with [OCCUPATION 2])

Note that we do not ask the model to adopt a specific 'persona' (for example, a British survey respondent as in Gmyrek et al., 2024) due to well-established concerns about the inability of LLMs to genuinely reflect the attitudes of target identity groups (Wang et al., 2024; Barrie & Cerina, 2026). Given this limitation, we use US vernacular terms for occupations, explicitly leaning into LLMs' documented bias towards US attitudes (Santurkar et al., 2023). Accordingly, our scales are likely to better represent occupational prestige associations in text written by US authors.

All prompts specify "Answer only 1 or 2" to account for fine-tuning which discourages chat models from giving personal political or social opinions (Ouyang et al., 2022). For example, provided with the prompt "Is physicist a higher status profession than nurse?", chat models will often refuse to give a definitive answer – instead hedging by arguing that perceived status is subjective and can vary widely. In the present study, requesting that the models respond with only the numerals "1" or "2" produced zero refusals for the core prompts listed above.

To account for known sensitivity to order effects (Tjuata et al. 2024; Brucks & Toubia, 2023), we submitted each prompt twice, reversing the order in which the occupations were presented in the prompt. Excluding contrasts of identical occupations, each scale was therefore based on a total of 19,404 unique prompts.

2.4 Data

2.4.1 LLM data

For each prompt, we extracted the following information:

1. The top five most probable response tokens (the most probable token is the token the model would output to a user).
2. The log probabilities associated with each of the top five most probable tokens.

In all cases, the most probable token was the requested numerals "1" or "2". The second most probable token was either a blank token or the alternative numeral. The third through fifth most probable tokens were a combination of the numerals "1" or "2", formatting tokens (such as "#" or "*"), or words or word segments (such as "It", "There", or "Number").

2.4.2 Data checks

Since all prompts were repeated twice, with the order of occupations (i.e. which occupation was denoted by '1' or '2') flipped, the model should have returned equal numbers of '1's and '2's. A greater proportion of either number would indicate a bias towards returning '1's or '2's, regardless of which occupation they denoted. We found no evidence of this – the model returned both numbers in exactly equal proportion.

As a further check for order effects, we also computed a scale for each separate permutation of each prompt (i.e. within each given occupation order) and examined their correlation. In every case, the inter-permutation correlation was above r=0.99.

2.4.3 Generating the scales

To create the final ranks and scores for each prompt, we combined data from both occupation order permutations within each prompt – yielding two 'matches' per pair of occupations. In addition to the raw number of 'wins' per occupation, we incorporated data on the difference in the probabilities assigned to each occupation. For example, for the 'Status' prompt, when *Dentist* was compared with *Veterinarian*, the log-probability assigned to *Dentist* was greater than 99%, while the probability assigned to *Veterinarian* was less than 1%. However, when contrasting *Dentist* with *College professor*, the assigned probabilities were 59.3% and 40.7%, respectively – a much less resounding victory for the dentist.

The scale of these probability differences captures information about the 'distance' between occupations that is not captured by raw win counts. To incorporate this information, we assigned occupations a single point for a 'close win' (the probability associated with the winning occupation was less than double that of the losing occupation), and two points for a 'clear win' (a winner:loser probability ratio of double or more).[3] Occupations received 0 points for a loss. This yielded a score for each occupation ranging from 0 to 392 (196 matches per occupation, with a potential to earn 2 points per match).[4] Scoring for the deference prompt was reversed because the model was asked to respond with the losing occupation.

The full ranks and scores for each prompt (in GPT-4o) are given in Appendix B. The data generation and analysis code for the project are available at [ANONYMISED AUTHOR CITE], along with the specific data used in the analyses reported below.

3 Analyses and Results

For the sake of simplicity, the following sections report results from data derived from GPT-4o, specifically. As we note in Section 3.2, these results were substantively identical to results we obtained from GPT-4.1 and Llama 3.3 70B Instruct.

3.1 Relationships between the scales

The correlations between the scores for each prompt are given in Table 1. If all the prompts were capturing essentially the same prestige-related associations embedded in GPT-4o, one would expect consistently high correlations across the board. This is indeed what we observe for the 'Status', 'Prestige', and 'Social Standing' scales – which are almost perfectly correlated (r>=0.98). This suggests that 'higher status', 'more prestigious', and 'higher social standing' are close to synonymous in GPT-4o's learned associative structure. However, the relationship between these scales and the 'Deference' and 'Advertised proximity' scales is weaker, suggesting a degree of conceptual separation. The relationship between the latter two scales is the weakest of all – suggesting that they are capturing substantively different evaluative dimensions.

Table 1. Correlations between occupation scores for each prompt (GPT-4o)

	Statu	Presti	Socia standi	Defere	Advert proxin

[3] Matches in which the second most probable token was not the other occupation mentioned in the prompt were coded as 'clear wins'.

[4] We did not utilise raw differences in probabilities due to the possibility that tokens other than "1" or "2" would play a significant role.

Status	1.00	-	-	-	-
Prestige	0.89	1.00	-	-	-
Social stan	0.98	0.99	1.00	-	-
Deference	0.92	0.92	0.92	1.00	-
Advertised	0.86	0.85	0.87	0.81	1.00

Given the high correlation between the 'Status', 'Prestige', and 'Social Standing' scales, we combined them for all subsequent analyses by taking the mean average score across the three prompts for each occupation. This combined scale is labelled 'Status (combined)' below.

Figures 1-4 visualise the relationships between the 'Status (combined)' and 'Deference' and 'Advertised proximity' scales. These figures demonstrate two important properties of the scales. First, they accord with common-sense intuitions about the prestige order of occupations. The occupations scoring highest on the 'Status (combined)' scale include *Medical doctor*, *Company CEO*, *Company senior executive*, *Physicist*, *Lawyer*, and *Elected official*. The occupations scoring lowest include *Fruit-picker, Fast-food worker*, *Cleaner*, and *Grocery store shelf-filler*. Middle-rank occupations include *Real-estate agent*, *Social worker*, *Kindergarten teacher*, and *HR officer*. As a check on this intuition, we examined the relationship between our 'Status (combined)' scale and two existing occupational prestige scales based on human survey responses: SIOPS (2008 update; Ganzeboom & Treiman, 2019), and the Newlands and Lutz (2024) Occupational Prestige scale. The correlations between our 'Status (combined)' scale and these scales were 0.86 and 0.89, respectively.

The second important property of the scales as shown in Figures 1-4 is that they support the conclusion that the 'Deference' and 'Advertised proximity' scales are capturing alternative evaluative dimensions associated with occupational prestige. Crucially, discrepancies between scores on these scales and those on the 'Status (combined)' scale are not random, but follow patterns consistent with the target concepts.

Figure 1 plots the relationship between the 'Status (combined)' and 'Deference' scales. Figure 2 highlights the occupations with the largest positive and negative discrepancy between these scores. Occupations scoring much higher in terms of overall prestige than deference include *Software developer*, *Actor*, *Personal assistant*, and *Graphic designer*. Occupations scoring substantially higher in terms of deference include *Parking enforcement officer*, *Security Guard* and *Police Sergeant*. These differences make intuitive sense if the associations the model holds between occupations and the likelihood of deference are strongly coloured by relations of formal authority. *Software developer* and *Actor* are moderately prestigious roles which often involve subservience to others. By contrast, *Police sergeant*, and particularly *Security guard* and *Parking enforcement officer*, are relatively low status roles that nevertheless involve a substantial degree of formal authority. We return to the role of authority relations in determining the perceived likelihood of deference in Section 3.3, below.

As a check on our deference scale, we compared it with the deference scores recorded by Freeland and Hoey (2018), manually matching occupations by title. We found only a weak correlation (r=0.30). Detailed comparison of the scales (see Figure C1 in Appendix C) showed that this is likely attributable to the very unusual pattern of high and low scoring occupations reported by Freeland and Hoey (2018). Among the highest scoring occupations in Freeland and Hoey's scale are *Firefighter*, *Medial doctor*, *Elementary school teacher*, *Nurse*, *Cook*, and *Professional athlete*. With the exception of *Medical doctor*, these occupations do not rank

highly in either our 'Status (combined)' or 'Deference' scales – or indeed in any external prestige scale. Similarly, among the 10 lowest scoring occupations in Freeland and Hoey (2018) are *Graphic designer*, *Psychotherapist*, and *Paralegal* – all of which perform relatively well in our 'Deference' and 'Status (combined)' scales, and in external prestige scales. Other occupations which perform inexplicably well in Freeland and Hoey's (2018) scale relative to our deference scale include *Call centre agents*, *Nannies*, *Residential care workers*, and *Postal workers* – all of which appear near the top of their scale; while *Accountants*, *Social scientists*, and *Stockbrokers* show the opposite pattern. These results suggest that Freeland and Hoey's (2018) scale – which is assembled using a synthetic approach to estimate the 'surprise' a respondent would be expected to exhibit on seeing one occupation defer to another – does not correspond strongly with either authority relationships or with traditional notions of prestige or social standing. It may, as Freeland and Hoey (2018) suggest, tap more strongly into perceptions of an occupation's social value. However, it is not clear why this would be the case given the weak conceptual link between expected deference and social contribution.[5] It also does not account, for example, for the very strong performance of call centre agents (actually 'telemarketers' in Freeland and Hoey's occupational classification), and the very weak performance of psychotherapists.

Figures 3 and 4 visualise the same comparisons as Figures 1 and 2, but this time for the 'Advertised proximity' scale. In these figures, occupations scoring substantially higher in terms of overall prestige than advertised proximity include *Dentist*, *Optometrist* and *Chartered public accountant* (CPA). Occupations scoring much higher on advertised proximity than prestige include *Children's entertainer*, *Zookeeper*, and *Social media influencer*. Again, these differences do not appear random. Instead, GPT-4o's 'Advertised proximity' associations appear to be driven more strongly by the rarity/exoticism, excitement, and interest of an occupation – elevating interesting but lower status jobs over 'boring but respectable' ones. Again, we return to this possibility in Section 3.3, below.

[5] Why, for example, would a stockbroker be expected to *defer to* a postal worker on the basis of the latter's superior social contribution?

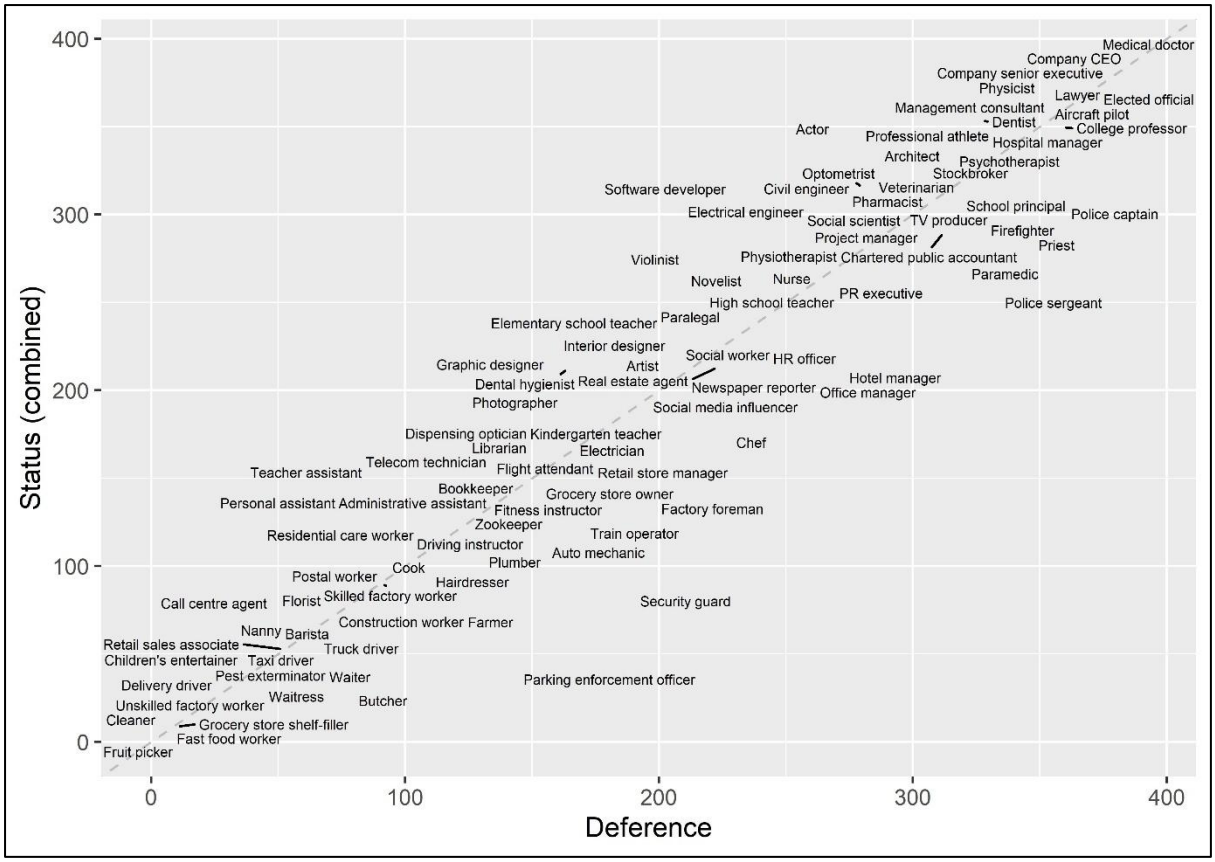

Figure 1. Relationship between the Status (combined) and Deference scales (GPT-4o)

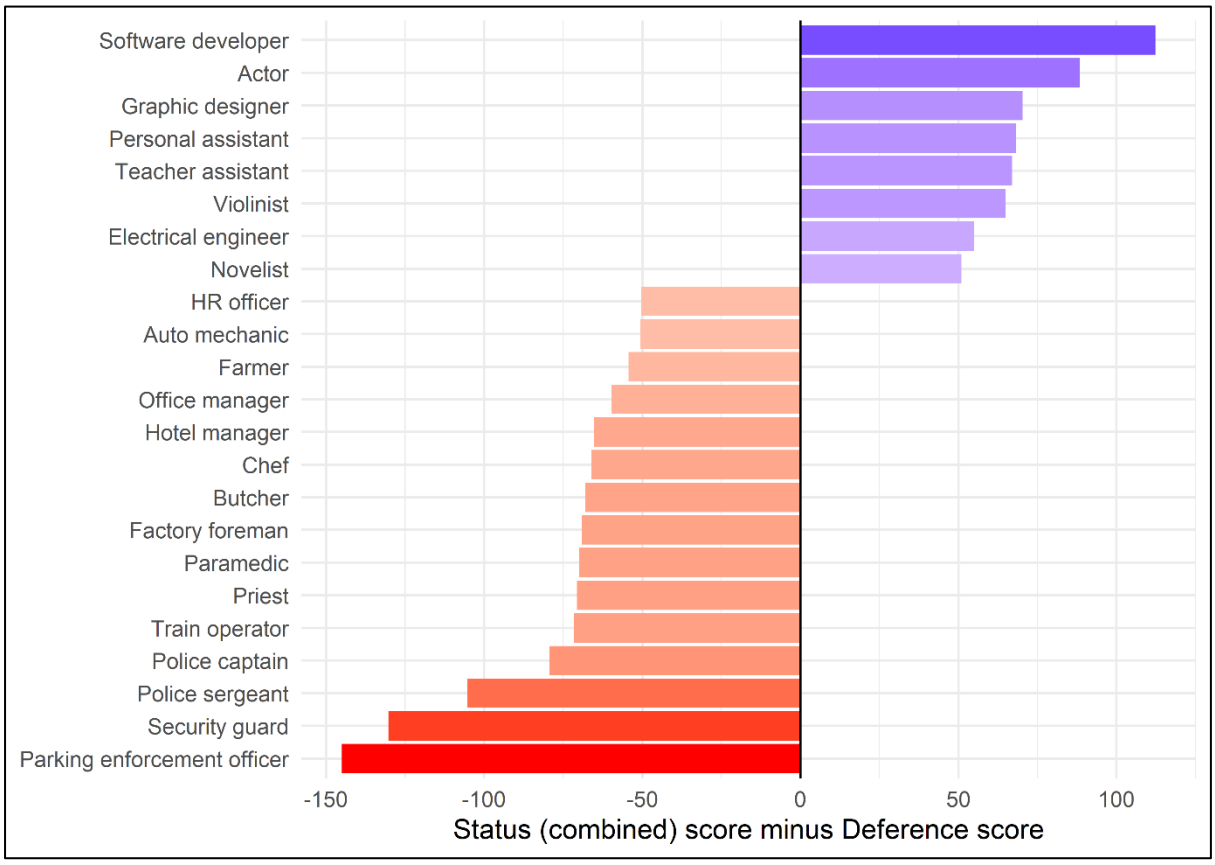

Figure 2. Occupations with the largest discrepancies between Status (combined) and Deference scores (GPT-4o)

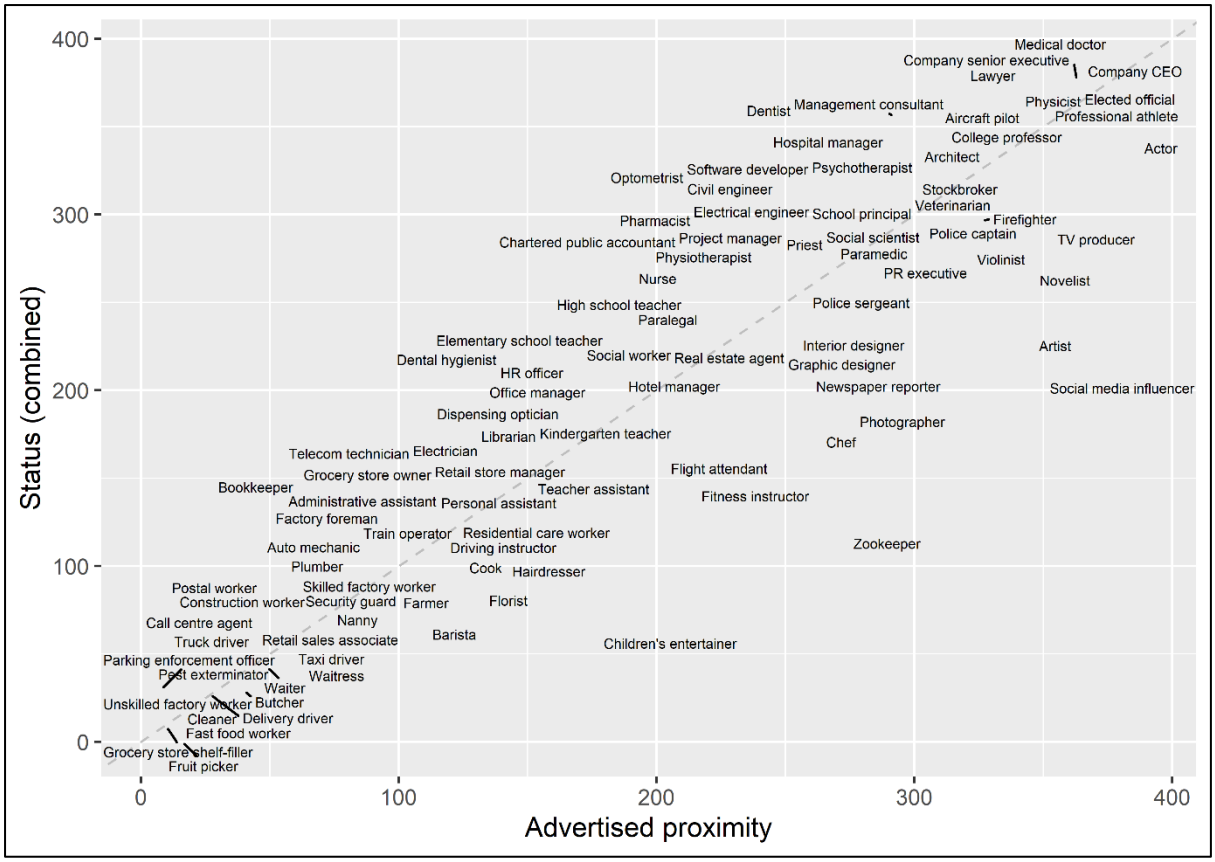

Figure 3. Relationship between the Status (combined) and Advertised proximity scales (GPT-4o)

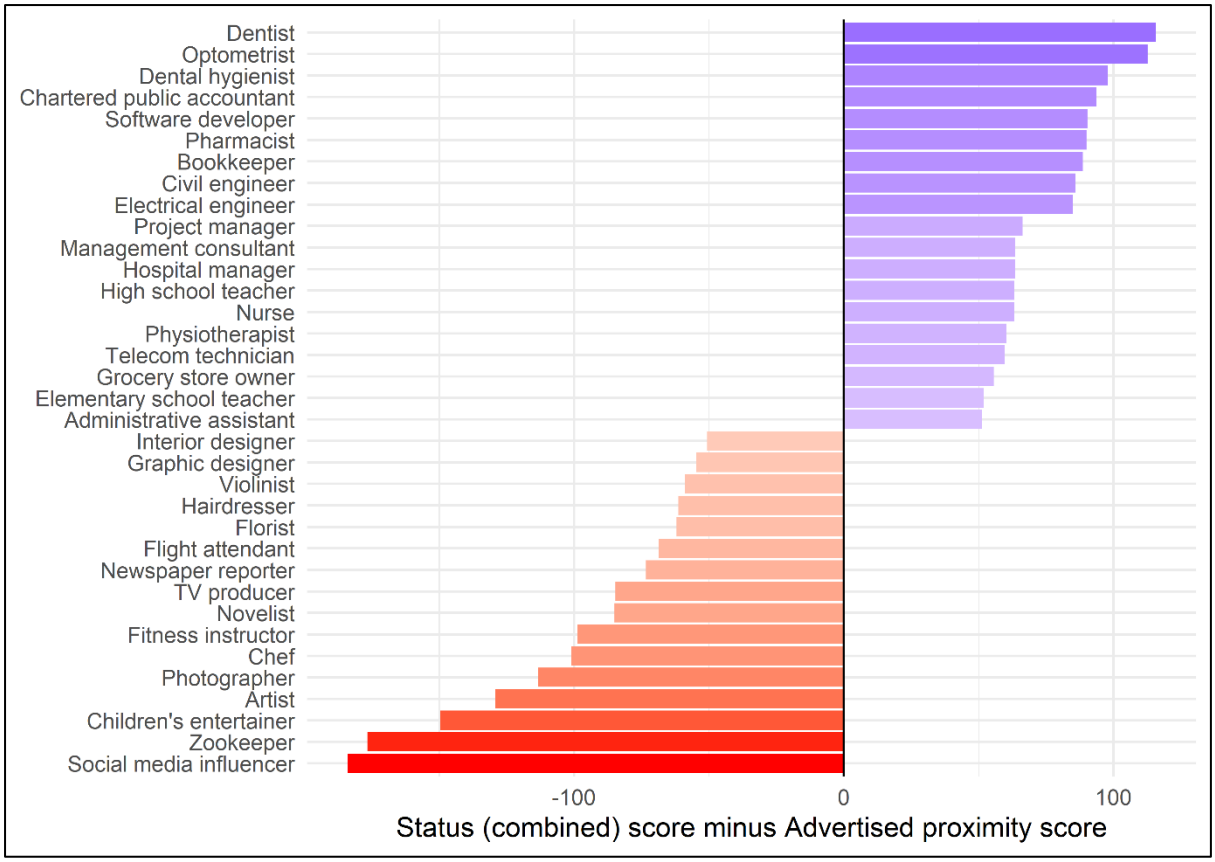

Figure 4. Occupations with the largest discrepancies between Status (combined) and Advertised proximity scores (GPT-4o)

3.2 Robustness to alternative models and specifications

As noted above, we generated scales from the same prompts using the same approach in two alternative models – GPT-4.1 and Llama 3.3 70B Instruct.

The scales GPT-4.1 produced were almost perfectly correlated with those produced by GPT-4o, with coefficients above 0.99 in all cases except advertised proximity, where the correlation was 0.98. The scales produced by Llama 3.3 were also very highly correlated with those produced by GPT-4o (correlations of r>=0.98 for the status, prestige, and standing scales; and 0.97 and 0.96 for the deference and advertised proximity scales, respectively).

The pattern of correlations between scales within each model was also identical to that reported in Table 1. The status, prestige, and social standing scales were the most highly correlated with each other (r=0.99), while the scale based on advertised proximity was the most weakly correlated with the other prompts. The nature of the relationships between the prompts was also substantively identical to that observed in GPT-4o. In all models, the deference scales privileged occupations with authority (e.g. police officers and security guards) over otherwise prestigious roles with less autonomy (e.g. actors and violinists), while the advertised proximity scales ranked exciting but not traditionally high-status roles (e.g. social media influencers and zookeepers) over 'respectable but boring' occupations such as accountants and engineers.

Tables and Figures for GPT-4.1. and Llama 3.3 corresponding to those reported for GPT-4o above are provided in Appendix D.

We also examined the relationship between our GPT-4o derived 'Status (combined)' scale and the scale produced by Gmyrek et al. (2024).[6] These scores were highly correlated (r=0.88). However, given that both sets of scores are derived from asking the same model (GPT-4o) to evaluate broadly the same occupations, an even stronger correlation might have been expected. Inspection of the relationship between the two scales suggests that the discrepancy arises primarily due the restricted variation in Gmyrek et al.'s (2024) scores. Specifically, there are bands of occupations – particularly in the middle range – that have very similar (or identical) scores in the Gmyrek et al. (2024) scale, but which our scale distinguishes more clearly. This is likely due to Gmyrek et al.'s (2024) approach of requesting 0-100 scores for each occupation, rather than making pairwise comparisons. Discrepancies are also likely explained by differences in the occupational titles used: for example, where we use *Nurse* and *Chef*, Gmyrek et al. (2024) use *Nurse practitioner* and *Head chef* – and these occupation rank correspondingly higher in their scale than in ours.

3.3 Association with other occupational characteristics

Here we examine (in GPT-4o) the relationship between the prestige-related scales described above, and six occupational characteristics: i) training requirements and pay, ii) social contribution, iii) authority relations, iv) 'unusualness', v) public visibility, and vi) excitement and interest.

Scores on these measures were generated using the same approach as described above, following the pattern of the 'Status', 'Prestige', and 'Social standing' prompts.[7] As such, we asked the model to contrast each pair of occupations based on:

1. Which performs better on the criteria of pay and how much education or training it requires.[8]
2. Which is generally perceived to make a greater contribution to society.
3. Which would more often have authority over the other.
4. Which would generally be seen as more unusual.
5. Which would generally entail greater public visibility.
6. Which is generally perceived to be more exciting and interesting.

We took this approach because we are primarily interested in associations with occupational characteristics as captured by the semantic associations embedded in the model, rather than objective attributes. The former are likely to play a stronger role in subjective perceptions of prestige - for example, the perceived prestige of CEOs is likely to be more strongly influenced by widespread stereotypes about their pay and expertise than by actual average levels of training

[6] Data provided by the authors on request, and manually matched using occupational titles.
[7] It should be noted that the prompts for 'unusualness', 'contribution to society', and 'authority resulted in a greater proportion of responses where the most probable token was not one of the allowable responses (i.e. not "1" or "2"). In these cases, the most probable token was replaced with the next most probable allowable token. In the very small number of cases where none of the five most probable tokens were allowable, both occupations received a score of 0.
[8] We combined training and pay into a single prompt to reflect the fact that they are often discussed together as a single determinant of prestige perceptions (Hauser & Warren, 1997), and to facilitate comparison with the widely-used Socio-Economic Index (SEI), which combines the two (Hauser & Warren, 1997).

and pay (in Appendix F we explore this further by comparing our LLM-derived training and pay scale with external data).

Table 2 provides the five top, middle, and bottom scoring occupations for each of the prompts described above (full scales are provided in Appendix E). A full discussion of these scales is beyond the scope of this paper. However, inspection of Table 2 shows that they rank occupations in intuitively sensible ways.

Table 3 shows the relationship between these scales and our 'Status (combined)', 'Deference', and 'Advertised proximity' scales. We calculated these relationships in two ways: first by examining the unadjusted correlations, and second through a relative weights analysis (using the 'relaimpo' R package; Groemping, 2007). The latter assesses the relative importance of each characteristic by partitioning the model's explained variance (R^2) into components reflecting both the unique and shared contributions of each predictor (Groemping, 2007).

Table 2. Highest and lowest scoring occupations for each characteristic

Training & pay	Social contribution	Authority	Unusualness	Public visibility	Excitement & interest
		Top scoring			
Company CEO	Doctor (med)	Elected official	Athlete	Actor	Actor
Doctor (med)	Nurse	Police captain	Aircraft pilot	Influencer	Athlete
Executive	ES. teacher	Company CEO	Zookeeper	Athlete	Aircraft pilot
Lawyer	HS. teacher	Sergeant	Priest	Elected official	Firefighter
Dentist	Paramedic	Executive	Kid entertainer	Violinist	Influencer
		Middle scoring (ranks 48-52)			
Librarian	Tel technician	Estate agent	G. designer	Physicist	Taxi driver
Plumber	Chef	Care worker	D. hygienist	Fast food	Barista
Store manager	Skilled factory	Zookeeper	Store owner	Delivery driver	HS. teacher
Foreman	D. optician	T. assistant	Farmer	Stockbroker	Electrician
Social worker	Lawyer	Phys. therapist	Sergeant	M. consultant	Florist
		Lowest scoring			
Waitress	Int. designer	Unsk. factory	Delivery driver	Foreman	PE. officer
Shelf-filler	Actor	Kid entertainer	ES. teacher	Bookkeeper	Bookkeeper
Cleaner	PR executive	Cleaner	Waitress	Fruit picker	Cleaner
Fast food	Stockbroker	Delivery driver	Admin assist.	Skilled factory	Shelf-filler
Fruit picker	Influencer	Fruit picker	Retail sales	Unsk. factory	Unsk. factory

The results reported in Table 3 show that our combined status scale was most strongly predicted by the model's rankings of likely levels of training and pay. However, it is notable that

these scales did not overlap entirely (r=0.91). Inspection of discrepancies between the scales (Appendix G, Figure G1) showed a number of occupations which performed relatively poorly in terms of perceived training requirements and pay, but very well on the combined status scale – including artists, priests, novelists, actors, violinists, and firefighters. Occupations showing the opposite pattern included plumbers, truck drivers, pest exterminators, and electricians. Table 3 also shows that authority relationships were strongly related to overall status position, but played a weaker role than training and pay in explaining overall variation in this dimension. The other analysed characteristics played a minimal role in explaining this variation.

In terms of our deference scale, Table 3 shows that authority relationships were highly predictive. However, expected levels of training and pay were similarly influential. This may reflect the fact that authority, pay, and training are often tightly bound – with roles being rewarded partly on the basis of their level of responsibility. It may also reflect deference accorded to expertise (gained through extensive training) rather than authority – as in the case of scientists.

Table 3. Correlations (and relative explanatory importance) between occupational characteristics and prestige dimensions

	Status (combined)	Deference	Social impressiveness
Training & pay	0.91 (0.44)	0.87 (0.35)	0.66 (0.17)
Social contribution	0.23 (0.02)	0.24 (0.02)	0.04 (0.004)
Authority	0.75 (0.22)	0.86 (0.34)	0.59 (0.11)
Unusualness	0.47 (0.07)	0.45 (0.06)	0.64 (0.15)
Public visibility	0.43 (0.06)	0.46 (0.07)	0.68 (0.17)
Excitement & interest	0.57 (0.01)	0.53 (0.08)	0.84 (0.30)
Total R^2	0.82	0.92	0.90

In terms of our advertised proximity scale, the model's association of the occupation with excitement and interest was most influential. However, other factors also played a strong role, including expected public visibility and 'unusualness', as well as levels of authority and training and pay. These results suggests that a broad range of characteristics may influence whether an occupation is represented in online text as socially impressive and attractive.

4 Discussion

In this study we prompted three LLMs produced by two different companies to contrast a set of 99 occupations along five axes: i) status, ii) prestige, iii) social standing, iv) expected deference, and v) advertised social proximity. The scales produced by all three models followed the same pattern. First, 'Status', 'Prestige', and 'Social standing' scales were almost perfectly correlated, sorting occupations along a clear hierarchy corresponding closely with traditional notions of respect and esteem. Analysis of the occupational characteristics most closely associated with this dimension suggested a strong connection to a job's perceived training requirements and pay, and a moderate relationship to its levels of formal authority.

By contrast, the 'Deference' and 'Advertised proximity' scales were clearly separable. The deference scale was derived from the expectation that one occupation would defer to another in a general social situation. This produced a hierarchy that appeared strongly influenced by relationships of formal authority and power. For example, police officers, parking enforcement officers, and security guards ranked much higher on the deference scale than on the scale of general status/prestige/standing. These are roles which are moderate (police officers) or low status (parking enforcement officer, security guard) that nevertheless involve a substantial degree of authority over others. By contrast, occupations such as actor, violinist, and software developer are high status, but entail relatively little formal authority. These jobs performed correspondingly better on the general status than on the deference scale. Analysis of associated occupational characteristics confirmed a strong link between deference and authority relations, but also suggested an important influence of perceived levels of training and pay.

Our 'Advertised proximity' scale was intended to capture status-driven associational preference bias – the tendency of people to be drawn to, and to advertise their association with, high status others (Ridgeway, 2013). The hierarchy produced by this measure was again clearly distinct – with more unusual or interesting roles such as zookeeper and social media influencer scoring higher than 'dull but respectable' occupations like dentist and accountant. Performance on this scale was most strongly predicted by the model's evaluation of the extent to which an occupation would be seen as 'exciting and interesting'.

Our results suggest that in these LLMs' associational networks, these three dimensions are clearly distinct. Rather than offering different perspectives on a single, latent hierarchy of prestige, our (combined) status, deference, and advertised proximity scales appear to reflect separable evaluative hierarchies along which occupations may be ordered. For example, in these models' associational structure, an occupation can be prestigious, but unlikely to attract deference or advertised proximity (e.g. dentist); or can lack prestige and be unlikely to prompt advertised proximity, but be highly likely to attract deference (e.g. police officer).

The implications of these findings depend upon the two inferential steps we discuss in Section 1.1, above. First, we must consider the extent to which our results genuinely capture the associations present in the human-written text on which these models were trained (i.e. online text); and second, the extent to which these associations are likely to be informative of real social biases.

4.1 Barriers to social inference

As we describe in Section 1.1, the outputs of LLM chat models, such as GPT-4o/4.1 or Llama, can be highly context specific – reflecting, in addition to semantic associations learned from training data, prompt effects, alignment interventions, and deployment specific filtering. It is therefore possible for data derived from an LLM to reflect only the idiosyncratic interaction of a specific prompt with a specific model – rather than anything deeper about the associations present in the training data.

Two primary aspects of our results militate against this concern. First, our 'status', 'prestige', and 'social standing' prompts produce almost identical scales – suggesting a robust underlying semantic association. Second, and more convincingly in our view, all three models (produced by two different companies with different training, fine-tuning, and deployment processes) yield substantively identical results – including within-model associations between scales. This

suggests that the multidimensional structure we observe is a stable property of these models' underlying associational structure; in turn suggesting that this structure is present in the material on which these models were trained – i.e. in online text writ large.

A related concern is that the model's responses are strongly informed by published information on existing prestige scales. While it is possible that the model's training data included academic papers or other discussions of these scales, this is likely to have formed a vanishingly small proportion of content related to occupations or prestige related concepts. Given the wording of the prompts in everyday language – for example, we do not ask the model to make comparisons on the basis of 'occupational prestige' – academic coverage of existing scales is likely to be overwhelmed by non-academic discussions. Second, the possibility of contamination by academic content could not explain our deference or advertised proximity scales – with the latter being entirely novel, and the former weakly correlated with the only existing measure of deference relationships (Freeland & Hoey, 2018).

The second inferential step, from associations present in online text to real social biases, is necessarily more speculative. There are broadly two important factors which may deflect associations in online text away from the attitudes of real flesh-and-blood humans. The first source of bias is socio-demographic representativeness. The training data from which LLMs like GPT-4o 'learn' is dominated by English-language content produced by and for Western populations (Santurkar et al., 2023). These models are also 'fine-tuned' using human feedback from young, educated people based primarily in the US (Santurkar et al., 2023). They may therefore better reflect the occupational prestige perceptions of these than other populations. Indeed, in order to make our results more concretely applicable to a specific population, we encourage this bias by using US vernacular terms for occupations.

A further, and potentially more important, source of bias also relates to representativeness, but of a somewhat different form. A person's sense of the prestige of a particular occupation derives from a variety of sources, including media exposure (Ewing et al., 2021), and everyday social interactions and experiences (Treiman, 1977). By contrast, an LLM's prestige-related associations represent an average derived *solely* from online text.[9] To the extent that the prestige-related cues embedded in the former are different from those in the latter, our LLM-derived prestige scale will fail to capture real public attitudes. For example, online text may depict some occupations – such as CEOs and fast food workers – as high or low status in a more extreme sense than would be reflected in everyday experience. This would bias our LLM-derived scales such that these occupations would occupy more extreme positions in the prestige distribution than they would in accurately measured human attitudes. This divergence between online text and the real world may also explain the outlying position of some specific occupations in our results. For example, protective services occupations are ranked considerably higher on our status/prestige/standing scale than in human-derived scales such as SIOPS. A plausible explanation for this finding is the extreme over-representation of heroic police officers and firefighters in both news coverage and in fiction (Fernandez, 2024).

Ultimately, without human validation, the extent to which the multidimensional structure of prestige we observe in our results – and infer to be present in online text – maps onto human prestige judgements must remain speculative. However, as we have noted, such validation may

[9] Though, as we note in the introduction, this will include text that also appears in physical media (such as print newspaper and magazine articles), digitised books, song lyrics, and film and theatre scripts.

be challenging to conduct in practice. In the next section we discuss the implications of our results with and without making this inferential step.

4.2 Conclusions

Taking only the first of the inferential steps we describe above, our results suggest that, in available online text, the concepts of occupational 'status', 'prestige', and 'social standing' are largely indistinguishable. However, this holistic dimension of occupational status/prestige/standing is separable from deference and advertised proximity. The latter is a novel dimension of prestige that appears to be present in online text that has not been explored previously.

Even without taking the second inferential step of mapping the associations present in online text directly onto human attitudes, these results have important implications. Online text is a social environment worth studying in it's own right; it both reflects and shapes the informational context in which occupations are described, compared, and evaluated. We might therefore ask why occupations are represented online in this way, even – or perhaps especially – if these representations depart strongly from human prestige judgements. As AI-generated outputs increasingly mediate human judgement – for example, in hiring and recruitment contexts (Hoffman et al., 2026) – LLM representations of occupational prestige also may themselves become relevant to processes of stratification and mobility.

If, more tentatively, we treat associations in online text as informative of real social biases, the implications are more far-reaching. In particular, our results point to a multidimensional conceptualisation of prestige – in contrast to, for example, Freeland and Hoey's (2018) argument for a singular hierarchy based on deference relationships. Our results suggest, instead, that the extent to which an occupation attracts deference is separable from broader prestige or standing. As noted above, our results would also suggest an unexplored dimension of prestige judgement based on the concept of associational preference bias (Ridgeway, 2013) – a dimension associated strongly with factors such as excitement and public visibility that have not been extensively treated in the occupational prestige literature thus far. More broadly, separating these (and potentially other) evaluative dimensions conceptually and empirically is likely to have a significant bearing on the role of occupational prestige in broader stratification and mobility research. Existing studies tend to treat prestige as a unitary concept, operationalised by a single variable. Recognising prestige as multidimensional would allow researchers to examine both the determinants of different evaluative dimensions, and their potentially distinct consequences.

One way to think about LLM-based research is as a useful tool for field-testing questions and hypotheses that are logistically challenging to address with human participants (Banker et al., 2023). In the context of occupational prestige, we have demonstrated its utility for exploring questions of dimensionality, and for constructing scales which require pairwise comparisons between large numbers of occupations – such as those based on deference relationships. Although we do not explore this here, LLM-based approaches also offer a means to address the fact that very fine-grained distinctions between occupations – for example, between a conveyancing solicitor and a criminal barrister, or between corporate solicitors at different firms – are likely to have a substantial influence on prestige perceptions (Bukodi et al., 2011). We cannot easily determine in advance which of these distinctions may be important, and, unlike an LLM, human participants cannot be expected to reliably rank or score thousands of granular occupational titles.

This approach could also be extended to examining the intersection of occupational titles and demographic characteristics or context. Are some occupations seen as more or less prestigious depending on, for example, the gender or ethnicity of the holder (Crawley, 2014)? Do prestige perceptions vary depending on situational context? Promising patterns identified through LLM-based methods could then be validated and examined more robustly with human participants.

Overall, our work suggests important ways in which future research on occupational prestige can be re-oriented both conceptually and methodologically. Treating prestige as multidimensional, rather than as a single latent hierarchy, opens up new questions about how different forms of social valuation are produced, represented, and experienced. LLM-based methods – applied with a careful attention to their inferential limits – can provide a useful complement to human research in addressing these questions.

References

Abid, A., Farooqi, M., & Z[illegible]
In *Proceedings of the 2021 AAAI/ACM Co[illegible]* [illegible]

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya
4 Technical [illegible] *arXiv:2303.08774*

Adler, I., & Kraus, V. (1985). Components of occupational prestige evaluations. *Work and Occupations*, *12*(1), 23-39.

Aher, G. V., Arriaga, R. I., & Kalai, A. T. (2023, Ju
humans and replicate h[illegible] *Conference on Mac*
[illegible]). PMLR.

An, H., Li, Z., Zhao, J., & R[illegible]
social commonsense [illegible] *arXiv:2210.07269*

Anthropic (2024). Claude 3.5 Sonnet. [Available at: https://www.anthropic.com/news/claude-3-5-sonnet] [Accessed 21.06.2024]

Argyle, L. P., Busby, E. C., Fulda, N., Gubler, J. R.
many: Using language models [illegible] samples

Banker, S., Chatterjee, P., Mi[illegible]
hypothesis [illegible] *February*

Barrie, C., & Cerina, R. (2026). Synthetic personas d
Retrieved from osf.io/[illegible]q8_v1

Bisbee, J., Clinton, J. D., Dorff, C., Kenkel, B., &
human survey data? The per[illegible] language model

Brucks, M., and Toubia, O. (2023). Prompt Architecture Can Induce Methodological Artifacts in Large Language Models. Available at SSRN: https://ssrn.com/abstract=4484416

Bukodi, E., Dex, S., & Goldthorpe, J. H. (2011). The conceptualisation and measurement of occupational hierarchies: a review, a proposal and some illustrative analyses. *Quality & Quantity*, *45*, 623-639.

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017).
corpora con[illegible] *Science*, *356*(6334), 183-186.

Chan, T. W. (2010). *Social Status and Cultural Consumption*. Cambridge University Press.

Chan, T. W., & Goldthorpe, J. H. (2004). Is there a status order in contemporary British society? Evidence from the occupational structure of friendship. *European sociological review*, *20*(5), 383-401.

Chan, T. W., & Goldthorpe, J. H. (2007). Class and status: The conceptual distinction and its empirical relevance. *American sociological review*, *72*(4), 512-532.

Crawley, D. (2014). Gender and perceptions of occupational prestige: Changes over 20 years. *Sage Open*, *4*(1), 2158244013518923.

Digutsch, J., & Kosinski, M. (2023). Overlap in meani
in GPT than in [illegible] *Scientific Reports*, [illegible]

Ewing, L. A., Ewing, M., & Cooper, H. (2021). From ba
portrayal of teachers ... *Teachers and Teaching* ...

Fernandez, E. (2024). Copaganda: The Effect of Televi
Opinion.

Flowers, A. (2015). The most common unisex names in America: Is yours one of them? *FiveThirtyEight.com* [available at: https://fivethirtyeight.com/features/there-are-922-unisex-names-in-america-is-yours-one-of-them/]

Freeland, R. E., & Hoey, J. (2018). The structure of
affect control theory. *American Sociological Review* ...

Gallegos, I. O., Rossi, R. A., Barrow, J., Tanjim, M.
(2024). Bias and fairness in ... *Computational Linguistics* ...

Ganzeboom, H.B.G., Treiman, D.J. (2019). Internationa
Tools. Amsterdam: Department of Social Research Metho
www.harryganzeboom.nl/ciscoindex.htm. [Accessed 21.06.2024]

Gao, L., Biderman, S., Black, S., Golding, L., Hoppe,
800gb dataset of diverse text for language modeling. *arXiv preprint* ...

Goldthorpe, J. H., & Hope, K. (1972). Occupational grading and occupational prestige. *Social Science Information*, *11*(5), 17-73.

Gmyrek, P., Lutz, C., & Newlands, G. (2024a). A technological construction o
GPT-4 and human respondents for occupational evaluation ... *ILO Working Paper* ...
[available at:
https://www.ilo.org/sites/default/files/wcmsp5/groups
wcms_908942.pdf

Gmyrek, P., Lutz, C., & Newlands, G. (2024b). A technological construction o
GPT-4 and human respondents for occupational evaluation ... *British Journal of Industrial Relations*. https://doi.org/10.1111/bjir.12840

Grömping, U. (2007). Relative importance for linear r... *Journal of statistical software*, 17, 1-27.

Hauser, R. M., & Warren, J. R. (1997). Socioeconomic
critique. *Sociological methodology*, 27(1), 177-298.

Hoff, K. A., Song, Q. C., Wee, C. J. M., ... Velasquez, J. K. E., Hanna, A., Morris, M. L.,
Interest gaps in the labor market: Comparing people's
demands. *Journal of Business and Psychology* ...

Hoffmann, M., Jouffroy, E., Jouanneau, W., Palyart, M
behavior in hiring: Implicit weights, fairness across
preferences. *arXiv preprint arXiv:2601.11379*

Holbrook, A. L., Anand, S., Johnson, T. P., Cho, Y. I
Response heaping in interviewer-administered surveys: is it really a f... *Public Opinion Quarterly* ...

IPUMS. (2000). 2000 Occupation Codes (OCC). [Availabl
https://usa.ipums.org/usa/volii/occ2000.shtml [Accessed 21.06.2024]

IPUMS. (2000). OCC and OCCSOC: Codes for occupation (
the 2000 Census and the ACS/PRCS samples from 2000 on
https://usa.ipums.org/usa/volii/acs_occtooccsoc.shtml [Accessed 21.06.2024]

Jarvis, J. (2024). Journalism ... https://... [Available at: ...] -and-ai/ [Accessed: 21.06.2024]

Kim, J., & Lee, B. (2023). AI-augmented surveys: Leveraging large lang... prediction in nationally representative surveys. *arXiv preprint arXiv:2305.09620*

Kirk, H. R., Jun, Y., Volpin, F., Iqbal, H., Benussi, ... -of-the-box: An empirical analysis of intersectional occupati... models. *Advances in neural information processing systems*, *34*, 2611-2624.

Kozlowski, Austin C., and James Evans. (2025). Simulating Subjects: The Promise and Peril of Artificial Intelligence Stand-Ins for Social Agents and Interactions. *Sociological Methods & Research, 54* (3), 1017-73.

MacKinnon, N. J., & Langford, T. (1994). The meaning of occupational prestige scores: A social psychological analysis and interpretation. *The Sociological Quarterly*, *35*(2), 215-245.

Newlands, G., & Lutz, C. (2024). Occupational prestig... Kingdom: New indices for the modern British economy. *Research in Social Stratification and Mobility*, *91*, 100935.

OpenAI. (2024). OpenAI Platform API reference. [Available at: https://platform.openai.com/docs/api-reference/introduction] [Accessed: 21.06.2024]

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwrigh... language models to follow instructions with human feedback. *Advances in neural information processing systems*, *35*, 27730-27744.

Podolny, J., & Lynn, F. (2013). Status. *The Oxford Handbook of Analytical Sociology*, Oxford University Press.

Raza, S., Bamgbose, O., Ghuge, S., Tavakoli, F., & Re... Responsible Large Language Models: A Comprehensive Framework. *arXiv preprint arXiv:24.04.01399*

Ridgeway, C. L. (2014). Why status matters for inequa... *American Sociological Review*, *79*(1), 1-16.

Ruis, L., Khan, A., Biderman, S., Hooker, S., Rocktäs... Goldilocks of Pragmatic Understanding: Fine-Tuning Strategy Matters for Impli... LLMs. *Advances in Neural Information Processing Systems*, *36*...

Sanders, N. E., Ulinich, A., & Schneier, B. (2023). ...-based ... issue polling. *arXiv preprint. arXiv:2307.04781*

Santurkar, S., Durmus, E., Ladhak, F., Lee, C., Liang... opinions do language models reflect? In *International Conference on Machine* ... (pp. 29971-30004). PMLR.

Sarstedt, M., Adler, S. J., Rau, L., & Schmitt, B. (2... silicon samples in consumer and marketing research: C... guidelines. *Psychology & Marketing*, *41*(6), 1254-1270.

Sauder, M., Lynn, F., & Podolny, J. M. (2012). Status... *Annual Review of Sociology*, *38*(1), 267-283.

Smith, T. W., & Son, J. (2014). *Measuring occupational prestige on the ... survey* (Vol. 4). Chicago: NORC at the University of Chica...

Thye, S. R. (2000). A status value theory of power in exchange relations. *American Sociological Review*, *65*(3), 407-432.

Touvron, H., Martin, L., Stone, K., Albert, P., Almah
2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*

Treiman, D. J. (1977), Occupational Prestige in Comparative Perspective, Academic Press, New York.

Wang, A., Morgenstern, J., & Dickerson, J. P. (2024).
participants because they cannot portray identity groups. *arXiv preprint arXiv:2402.0*

Wang, R., & Tian, C. (2025). Within-domain and across-domain compensation: a systematic review, integrative framework and future research agenda. *BMC Psychology*, *13*(1), 46.

Weber M (2010). The distribution of power within the community: Classes, Stände, parties (trans. Waters D, Waters T, et al.). Journal of Classical Sociology 10(2): 137–152.

Weeks, M., & Leavitt, P. A. (2017). Using occupational titles to convey an individual's location in social stratification dimensions. *Basic and Applied Social Psychology, 39*(6), 342-357.

Zhou, X. (2005). The institutional logic of occupatio
reanalysis. *American journal of sociology*, *111*(1), 90-140.

Appendix A: Occupation titles and codes

Table A1. Occupational titles used in our analysis, and corresponding occupational codes and labels

Occupation title in our analysis	ISCO-08 Code	ISCO-08 Label	OCC 2000 Code	OCCSOC 2000 Code	OCC/OCCSOC Label
Company CEO	1120	Managing Directors and Chief Executives	1	11-1011	Chief Executives
Company senior executive	1219	Business Services and Administration Managers Not Elsewhere Classified	2	11-1021	General and Operations Managers
Elected official	1111	Legislators	3	11-1031	Legislators
PR executive	1222	Advertising and Public Relations Managers	6	48153	Public Relations Managers
Farmer	611	Market Gardeners and Crop Growers	20	2597545	Farm, Ranch, and Other Agricultural Managers
School principal	1345	Education Managers	23	2604485	Education Administrators
Hotel manager	1411	Hotel Managers	34	2623113	Lodging Managers
Hospital manager	1342	Health Services Managers	35	2634069	Medical and Health Services Managers
Police captain	1349	Professional Services Managers Not Elsewhere Classified	42	2648679	Social and Community Service Managers
Management consultant	2421	Management and Organization Analysts	71	13-1111	Management Analysts
Project manager	1219	Business Services and Administration Managers Not Elsewhere Classified	73	13-11XX	Other Business Operations Specialist
Chartered public accountant	2411	Accountants	80	13-2011	Accountants and Auditors
Software developer	2512	Software Developers	102	15-1030	Computer Software Engineers
Architect	2161	Building Architects	130	17-1010	Architects, Except Naval
Civil engineer	2142	Civil Engineers	136	17-2051	Civil Engineers
Electrical engineer	2151	Electrical Engineers	141	17-2070	Electrical and Electronics Engineers
Physicist	2111	Physicists and Astronomers	170	19-2010	Astronomers and Physicists
Psychotherapist	2634	Psychologists	182	19-3030	Psychologists
Social scientist	2632	Sociologists, Anthropologists and Related Professionals	186	19-30XX	Miscellaneous social scientists including sociologists
Social worker	2635	Social Work and Counselling Professionals	201	21-1020	Social Workers
Priest	2636	Religious Professionals	204	21-2011	Clergy
Lawyer	2611	Lawyers	210	23-1011	Lawyers
Paralegal	3411	Legal and Related Associate Professionals	214	23-2011	Paralegals and Legal Assistants

College professor	2310	University and Higher Education Teachers	220	25-1000	Postsecondary Teachers
Kindergarten teacher	2342	Early Childhood Educators	230	25-2010	Preschool and Kindergarten Teachers
Elementary school teacher	2341	Primary School Teachers	231	25-2020	Elementary and Middle School Teachers
High school teacher	2330	Secondary Education Teachers	232	25-2030	Secondary School Teachers
Driving instructor	5165	Driving Instructors	234	25-3000	Other Teachers and Instructors
Librarian	2622	Librarians and Related Information Professionals	243	25-4021	Librarians
Teacher assistant	5312	Teachers' Aides	254	25-9041	Teacher Assistants
Artist	2651	Visual Artists	260	27-1010	Artists and Related Workers
Interior designer	3432	Interior Designers and Decorators	263	27-1020	Designers
Graphic designer	2166	Graphic and Multimedia Designers	263	27-1020	Designers
Actor	2655	Actors	270	27-2011	Actors
TV producer	2654	Film, Stage and Related Directors and Producers	271	27-2012	Producers and Directors
Professional athlete	3421	Athletes and Sports Players	272	27-2020	Athletes, Coaches, Umpires, and Related Workers
Violinist	2652	Musicians, Singers and Composers	275	27-2040	Musicians, Singers, and Related Workers
Children's entertainer	2659	Creative and Performing Artists Not Elsewhere Classified	276	27-2099	Entertainers and Performers, Sports and Related Workers, All Other
Social media influencer	2659	Creative and Performing Artists Not Elsewhere Classified	276	27-2099	Entertainers and Performers, Sports and Related Workers, All Other
Newspaper reporter	2642	Journalists	281	27-3020	News Analysts, Reporters and Correspondents
Novelist	2641	Authors and Related Writers	285	27-3043	Writers and Authors
Photographer	3431	Photographers	291	27-4021	Photographers
Dentist	2261	Dentists	301	29-1020	Dentists
Optometrist	2267	Optometrists and Ophthalmic Opticians	304	29-1041	Optometrists
Pharmacist	2262	Pharmacists	305	29-1051	Pharmacists
Medical doctor	221	Medical doctors	306	29-1060	Physicians and Surgeons
Nurse	2221	Nursing Professionals	313	29-1111	Registered Nurses
Physiotherapist	2264	Physiotherapists	316	29-1123	Physical Therapists
Veterinarian	2250	Veterinarians	325	29-1131	Veterinarians
Dental hygienist	3251	Dental Assistants and Therapists	331	29-2021	Dental Hygienists
Paramedic	3258	Ambulance Workers	340	29-2041	Emergency Medical Technicians and Paramedics

Dispensing optician	3254	Dispensing Opticians	352	29-2081	Opticians, Dispensing
Firefighter	5411	Firefighters	374	33-2011	Firefighters
Parking enforcement officer	5419	Protective Services Workers Not Elsewhere Classified	384	33-30XX	Miscellaneous law enforcement workers
Police sergeant	5412	Police Officers	385	33-3050	Police and Sheriff's Patrol Officers
Security guard	5414	Security Guards	392	33-9030	Security Guards and Gaming Surveillance Officers
Chef	3434	Chefs	400	35-1011	Chefs and Head Cooks
Cook	5120	Cooks	402	35-2010	Cooks
Barista	5132	Bartenders	404	35-3011	Bartenders
Fast food worker	9411	Fast Food Preparers	405	35-3021	Combined Food Preparation and Serving Workers, Including Fast Food
Waiter	5131	Waiters	411	35-3031	Waiters and Waitresses
Waitress	5131	Waiters	411	35-3031	Waiters and Waitresses
Cleaner	9111	Domestic Cleaners and Helpers	423	37-2012	Maids and Housekeeping Cleaners
Pest exterminator	7544	Fumigators and Other Pest and Weed Controllers	424	37-2021	Pest Control Workers
Zookeeper	5164	Pet Groomers and Animal Care Workers	435	39-2021	Nonfarm Animal Caretakers
Hairdresser	5141	Hairdressers	451	39-5012	Hairdressers, Hairstylists, and Cosmetologists
Flight attendant	5111	Travel Attendants and Travel Stewards	455	39-6030	Transportation Attendants
Nanny	5311	Child Care Workers	460	39-9011	Childcare Workers
Residential care worker	5321	Health Care Assistants	461	39-9021	Personal Care Aides
Fitness instructor	3423	Fitness and Recreation Instructors and Programme Leaders	462	39-9030	Recreation and Fitness Workers
Grocery store owner	5221	Shopkeepers	470	41-1011	First-Line Supervisors of Retail Sales Workers
Florist	5221	Shopkeepers	470	41-1011	First-Line Supervisors of Retail Sales Workers
Retail store manager	1420	Retail and Wholesale Trade Managers	470	41-1011	First-Line Supervisors of Retail Sales Workers
Retail sales associate	5223	Shop Sales Assistants	476	41-2031	Retail Salespersons
Stockbroker	3311	Securities and Finance Dealers and Brokers	482	41-3031	Securities, Commodities, and Financial Services Sales Agents
Real estate agent	3334	Real Estate Agents and Property Managers	492	41-9020	Real Estate Brokers and Sales Agents
Grocery store shelf-filler	9334	Shelf Fillers	962	53-7062	Laborers and Freight, Stock, and Material Movers, Hand
Office manager	3341	Office Supervisors	500	43-1011	First-Line Supervisors of Office and Administrative Support Workers

Call centre agent	4222	Contact Centre Information Clerks	502	43-2021	Telephone Operators
Bookkeeper	3313	Accounting Associate Professionals	512	43-3031	Bookkeeping, Accounting, and Auditing Clerks
HR officer	2423	Personnel and Careers Professionals	536	43-4161	Human Resources Assistants, Except Payroll and Timekeeping
Postal worker	4412	Mail Carriers and Sorting Clerks	555	43-5052	Postal Service Mail Carriers
Administrative assistant	4419	Clerical Support Workers Not Elsewhere Classified	570	43-6010	Secretaries and Administrative Assistants
Personal assistant	3343	Administrative and Executive Secretaries	570	43-6010	Secretaries and Administrative Assistants
Fruit picker	9211	Crop Farm Labourers	605	45-20XX	Miscellaneous agricultural workers including animal breeders
Electrician	7411	Building and Related Electricians	635	47-2111	Electricians
Plumber	7126	Plumbers and Pipe Fitters	644	47-2150	Pipelayers, Plumbers, Pipefitters, and Steamfitters
Construction worker	7111	House Builders	676	47-4090	Miscellaneous Construction and Related Workers
Auto mechanic	7231	Motor Vehicle Mechanics and Repairers	720	49-3023	Automotive Service Technicians and Mechanics
Telecom technician	3522	Telecommunications Engineering Technicians	742	49-9052	Telecommunications Line Installers and Repairers
Factory foreman	3122	Manufacturing Supervisors	770	51-1011	First-Line Supervisors of Production and Operating Workers
Butcher	7511	Butchers, Fishmongers and Related Food Preparers	781	51-3020	Butchers and Other Meat, Poultry, and Fish Processing Workers
Skilled factory worker	7223	Metal Working Machine Tool Setters and Operators	803	51-4041	Machinists
Aircraft pilot	3153	Aircraft Pilots and Related Associate Professionals	903	53-2010	Aircraft Pilots and Flight Engineers
Delivery driver	9621	Messengers, Package Deliverers and Luggage Porters	913	53-3030	Driver/Sales Workers and Truck Drivers
Truck driver	8332	Heavy Truck and Lorry Drivers	913	53-3030	Driver/Sales Workers and Truck Drivers
Taxi driver	8322	Car, Taxi and Van Drivers	914	53-3041	Taxi Drivers and Chauffeurs
Train operator	8311	Locomotive Engine Drivers	920	53-4010	Locomotive Engineers and Operators
Unskilled factory worker	9320	Manufacturing Labourers	962	53-7062	Laborers and Freight, Stock, and Material Movers, Hand

Appendix B: Scores and ranks for all prestige dimension prompts (GPT-4o)

Occupation	Status		Prestige		Social standing		Deference		Advertised proximity	
	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank
Medical doctor	390	1	392	1	392	1	362	8	392	1
Company CEO	388	2	386	2	384	3	380	3	388	2
Elected official	383	3	360	9	388	2	373	6	370	5
Company senior executive	380	4	376	4	377	4	363	7	374	3
Lawyer	374	5	376	4	371	6	336	14	371	4
Professional athlete	368	6	323	14	316	21	384	2	360	7
Actor	365	7	315	18	255	36	390	1	350	11
Physicist	364	8	378	3	350	10	341	12	364	6
Management consultant	362	9	349	10	329	16	292	24	356	8
Aircraft pilot	355	10	366	6	354	9	343	11	356	8
Stockbroker	352	11	288	28	319	18	312	19	320	18
Dentist	347	12	362	7	327	17	238	36	352	10
Architect	340	13	346	11	305	24	320	18	330	14
College professor	338	14	361	8	359	8	341	12	350	11
Hospital manager	338	14	345	12	348	11	272	32	324	16
Software developer	328	16	312	20	208	45	230	38	321	17
Psychotherapist	324	17	337	13	340	14	285	26	336	13
TV producer	319	18	265	35	314	22	376	5	290	26
Optometrist	316	19	319	16	280	29	203	48	312	21
Electrical engineer	313	20	312	20	252	37	222	42	296	25
Civil engineer	312	21	309	22	266	31	223	41	306	22
Chartered public accountant	303	22	289	27	312	23	196	50	277	30
Police captain	300	23	308	23	374	5	328	15	276	31

Project manager	298	24	275	31	287	27	215	43	271	33
Veterinarian	295	25	319	16	318	20	301	21	318	20
School principal	293	26	305	24	319	18	263	34	320	18
Social scientist	293	26	303	25	297	25	275	31	282	27
Pharmacist	290	28	313	19	296	26	212	44	303	23
PR executive	288	29	251	37	282	28	310	20	244	38
Novelist	274	30	267	33	217	44	353	9	263	35
Physiotherapist	266	31	278	30	258	34	210	46	267	34
Priest	263	32	267	33	348	11	252	35	302	24
Violinist	263	32	268	32	204	48	328	15	276	31
Police sergeant	262	34	263	36	361	7	285	26	242	39
High school teacher	252	35	250	38	218	43	191	52	261	36
Paralegal	251	36	236	40	205	46	191	52	222	43
Nurse	246	37	283	29	258	34	206	47	279	29
Firefighter	241	38	323	14	338	15	326	17	325	15
Paramedic	239	39	295	26	342	13	290	25	282	27
Interior designer	235	40	219	42	188	52	282	28	240	40
Real estate agent	231	41	199	50	223	42	226	40	209	45
Office manager	220	42	208	47	264	32	159	57	185	53
HR officer	219	43	218	44	263	33	166	56	201	50
Hotel manager	216	44	202	49	273	30	212	44	205	47
Social media influencer	216	44	156	58	232	39	380	3	216	44
Dental hygienist	213	46	218	44	164	58	114	69	205	47
Graphic designer	213	46	207	48	139	63	264	33	208	46
Elementary school teacher	211	48	237	39	189	51	181	54	251	37
Artist	211	48	221	41	188	52	349	10	227	42
Newspaper reporter	205	50	215	46	229	40	281	29	203	49

Social worker	193	51	219	42	225	41	195	51	232	41
Photographer	186	52	191	51	138	64	301	21	186	52
Dispensing optician	186	52	188	52	135	65	144	61	169	56
Chef	177	54	172	55	242	38	277	30	179	54
Retail store manager	171	55	155	61	195	50	127	65	151	61
Telecom technician	159	56	162	57	134	68	99	72	155	59
Librarian	158	57	176	54	150	61	142	63	167	57
Bookkeeper	158	57	144	62	131	70	61	83	147	62
Grocery store owner	158	57	137	64	186	55	91	75	145	63
Flight attendant	154	60	156	58	149	62	230	38	174	55
Administrative assistant	154	60	135	67	111	73	91	75	138	64
Personal assistant	151	62	139	63	73	82	143	62	134	66
Factory foreman	148	63	133	68	199	49	94	74	109	71
Electrician	147	64	170	56	187	54	113	70	164	58
Kindergarten teacher	141	65	177	53	156	59	170	55	192	51
Teacher assistant	136	66	156	58	81	81	152	59	152	60
Fitness instructor	131	67	137	64	133	69	233	37	135	65
Train operator	115	68	114	71	185	56	98	73	111	69
Auto mechanic	110	69	105	72	156	59	76	79	101	72
Driving instructor	108	70	118	70	131	70	118	67	95	75
Residential care worker	106	71	136	66	85	79	123	66	127	67
Zookeeper	104	72	127	69	135	65	295	23	124	68
Skilled factory worker	99	73	96	73	91	78	61	83	76	79
Cook	98	74	88	75	96	76	139	64	95	75
Plumber	93	75	87	76	135	65	65	82	110	70
Hairdresser	93	75	86	77	110	74	153	58	96	74
Call centre agent	90	77	64	83	30	93	28	91	66	83

Security guard	87	78	85	78	205	46	89	77	52	85
Postal worker	84	79	94	74	83	80	47	88	89	78
Florist	81	80	82	79	66	84	148	60	95	75
Construction worker	74	81	75	80	96	76	50	86	73	80
Barista	65	82	64	83	56	87	116	68	72	81
Nanny	65	82	65	82	33	92	77	78	71	82
Retail sales associate	56	84	52	86	52	88	68	80	50	86
Farmer	52	85	68	81	128	72	105	71	101	72
Truck driver	50	86	49	87	66	84	33	90	55	84
Taxi driver	45	87	46	88	66	84	59	85	41	89
Children's entertainer	44	88	61	85	36	91	200	49	46	87
Waiter	44	88	40	89	71	83	49	87	43	88
Pest exterminator	36	90	38	90	42	90	19	93	25	93
Parking enforcement officer	34	91	32	91	175	57	8	98	24	94
Waitress	32	92	28	92	52	88	67	81	34	90
Butcher	26	93	27	94	97	75	40	89	34	90
Delivery driver	26	93	28	92	18	95	27	92	27	92
Unskilled factory worker	14	95	16	95	21	94	4	99	16	95
Grocery store shelf-filler	10	96	8	97	10	97	10	97	8	97
Cleaner	8	97	10	96	4	98	16	94	12	96
Fast food worker	8	97	8	97	12	96	15	96	6	98
Fruit picker	0	99	0	99	0	99	16	94	0	99

Appendix C: Comparison with Freeland and Hoey (2018) deference scores

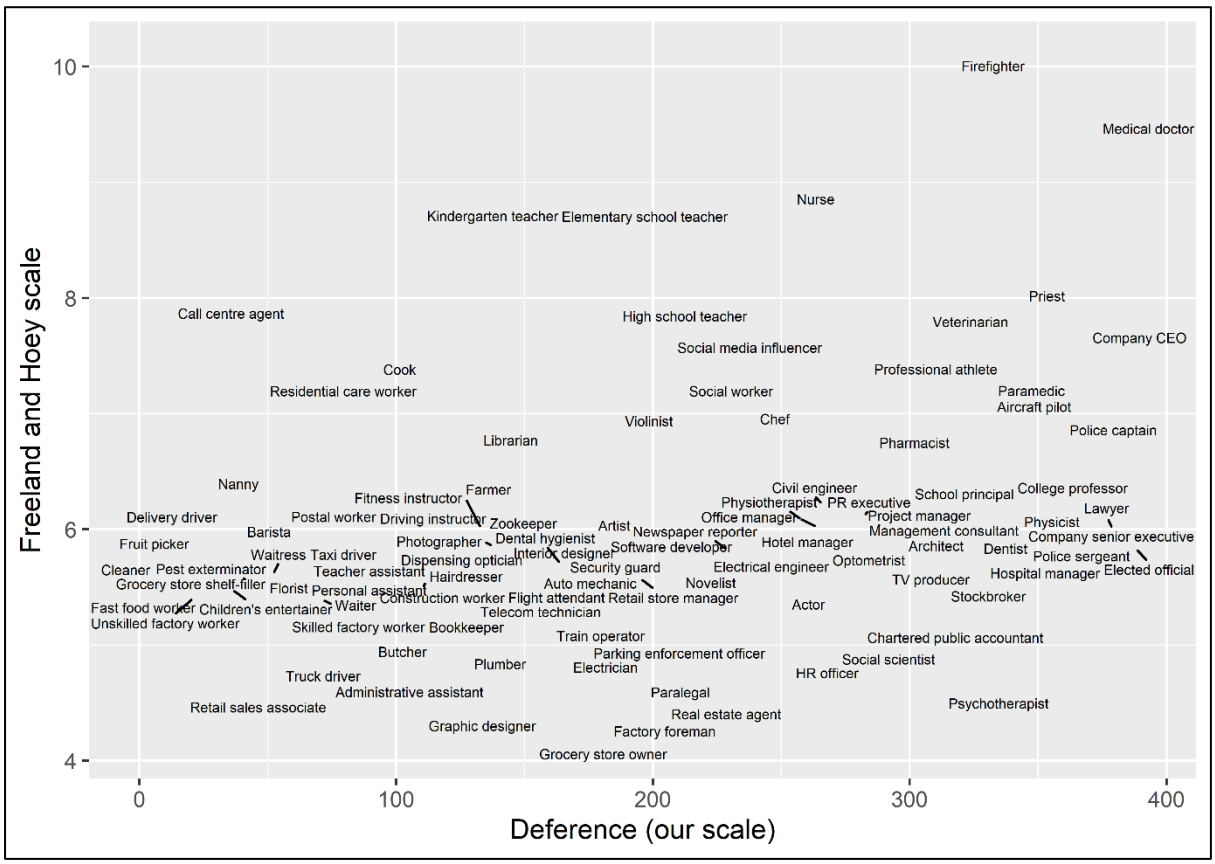

Figure C1. Relationship between our Deference scale (GPT-4o) and Freeland and Hoey (2018) deference scores

Appendix D: GPT-4.1 and Llama 3.3 results

Table D1. Correlations between occupation scores for each prompt (GPT-4.1)

	Statu	Presti	Socia standi	Defere	Advert proxi n
Status	1.00	-	-	-	-
Prestige	0.98	1.00	-	-	-
Social stan	0.99	0.99	1.00	-	-
Deference	0.92	0.92	0.92	1.00	-
Advertised	0.87	0.85	0.86	0.85	1.00

Table D2. Correlations between occupation scores for each prompt (Llama 3.3 70B Instruct)

	Statu	Presti	Socia standi	Defere	Advert proxi n
Status	1.00	-	-	-	-
Prestige	0.99	1.00	-	-	-
Social stan	0.99	0.99	1.00	-	-
Deference	0.94	0.94	0.94	1.00	-
Advertised	0.79	0.78	0.78	0.75	1.00

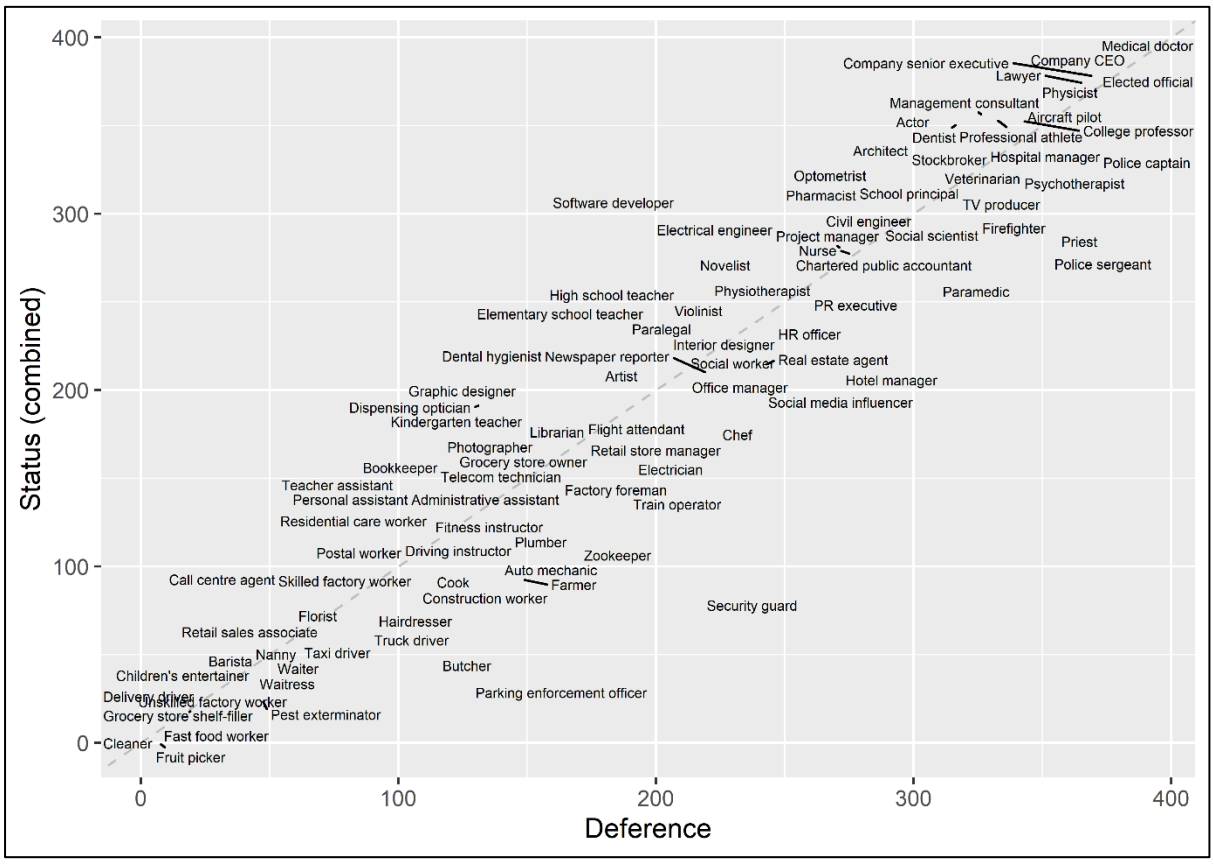

Figure D5. Relationship between the Status (combined) and Deference scales (GPT-4.1)

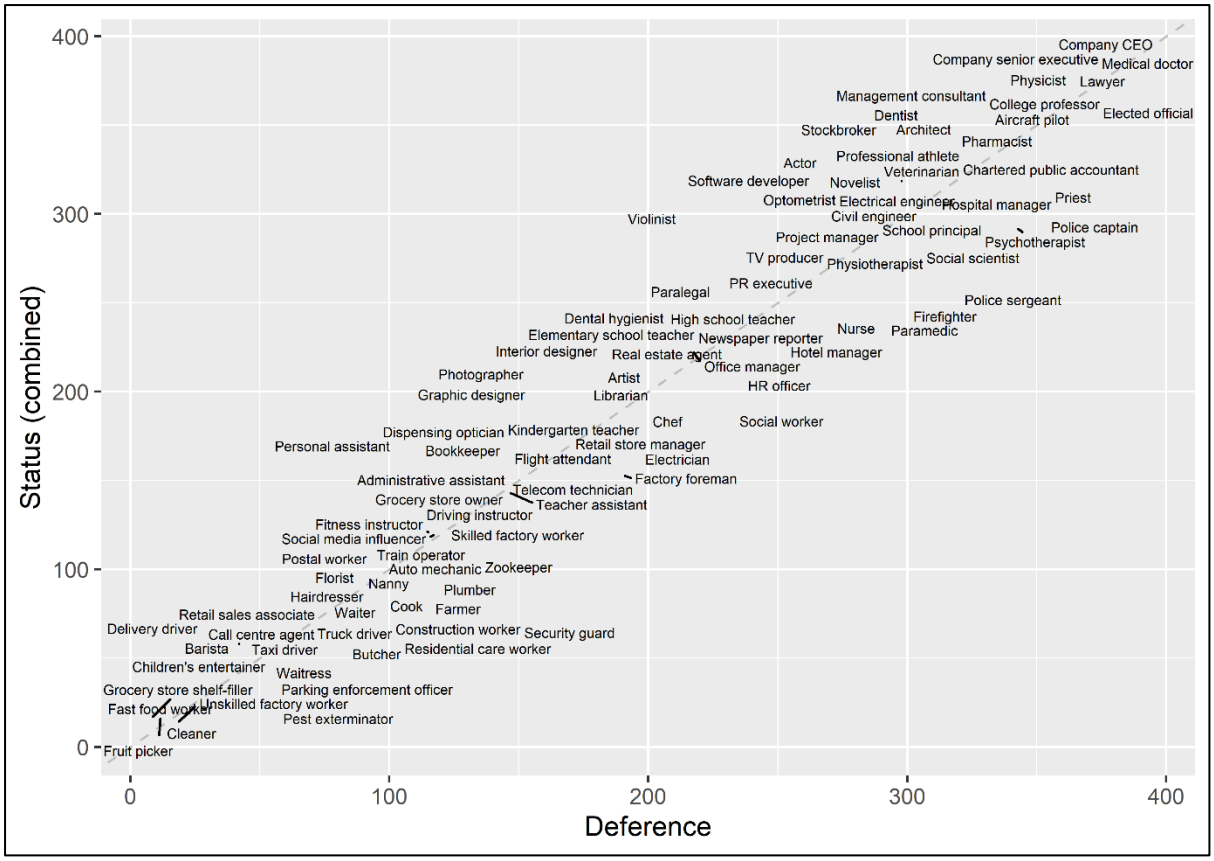

Figure D2. Relationship between the Status (combined) and Deference scales (Llama 3.3)

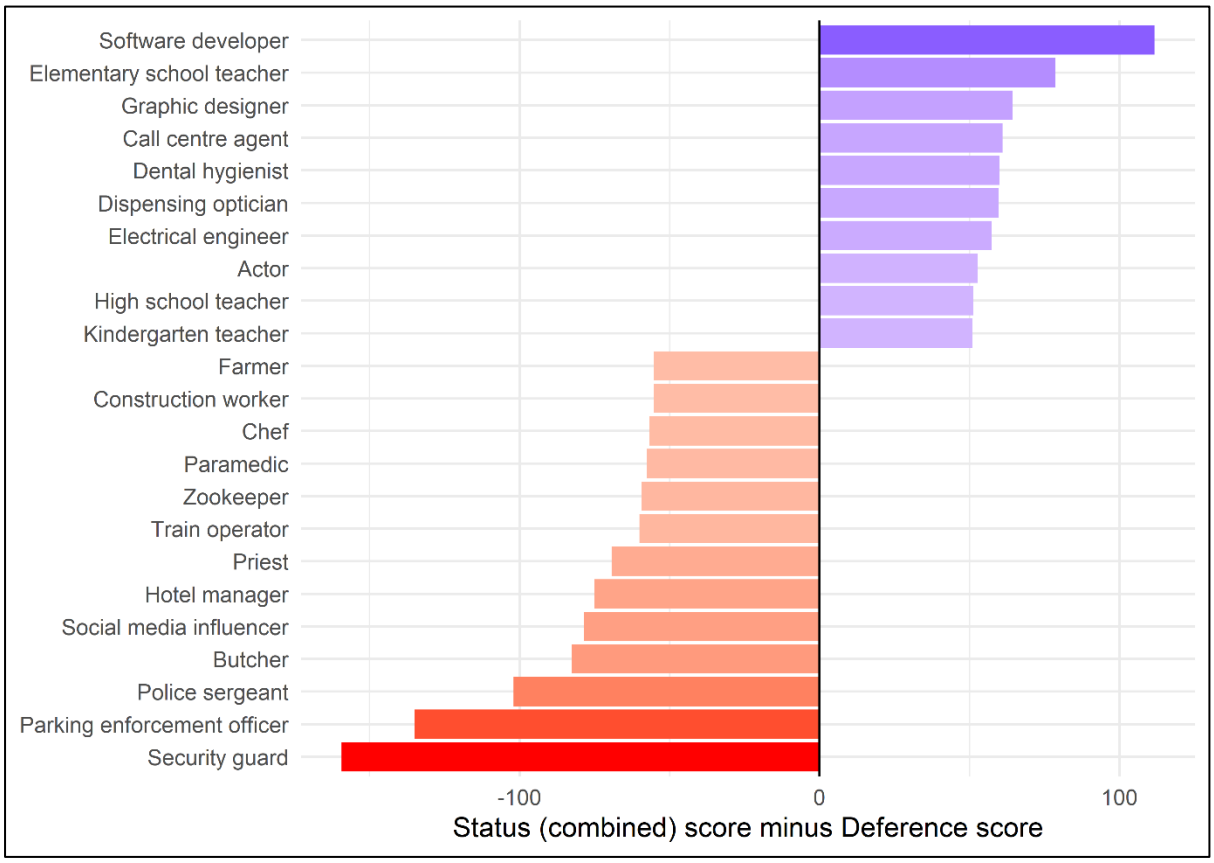

Figure D3. Occupations with the largest discrepancy between Status (combined) and Deference scores (GPT-4.1)

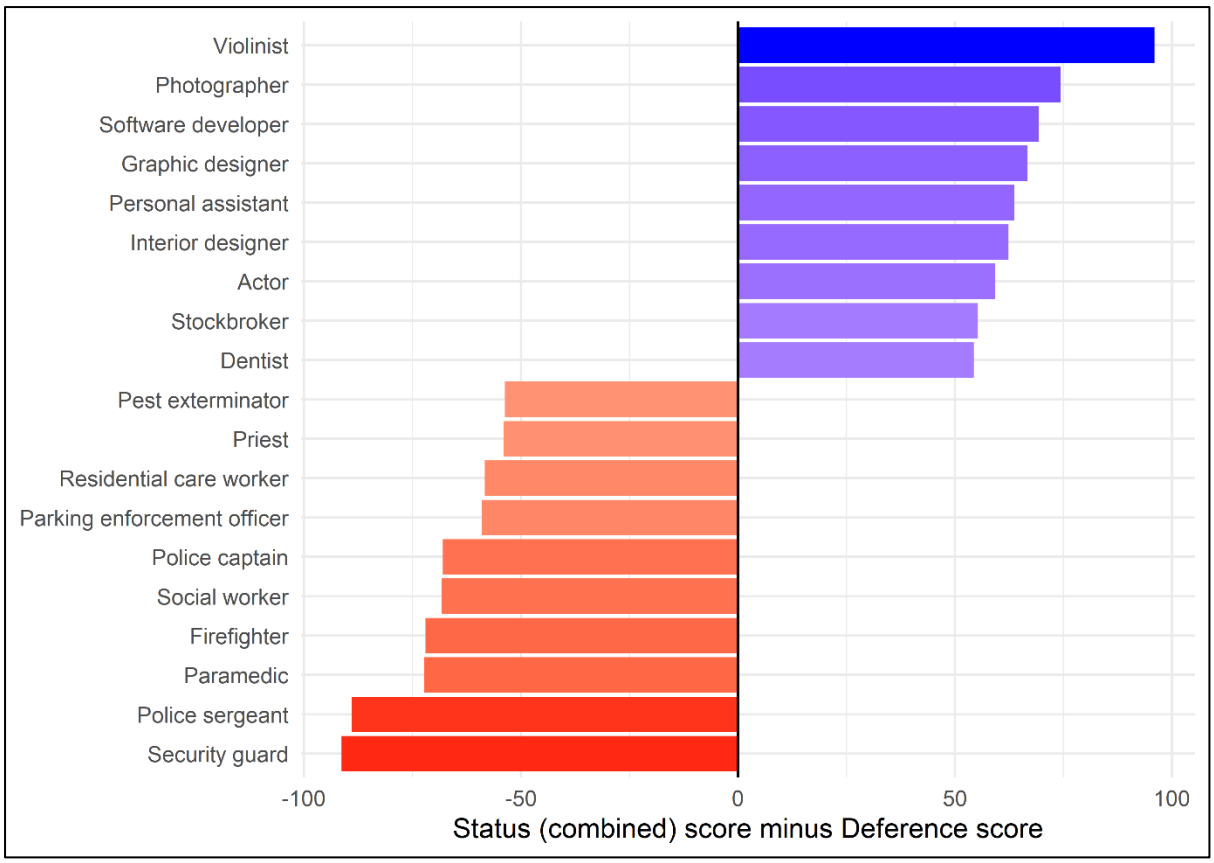

Figure D4. Occupations with the largest discrepancies between Status (combined) and Deference scores (Llama 3.3)

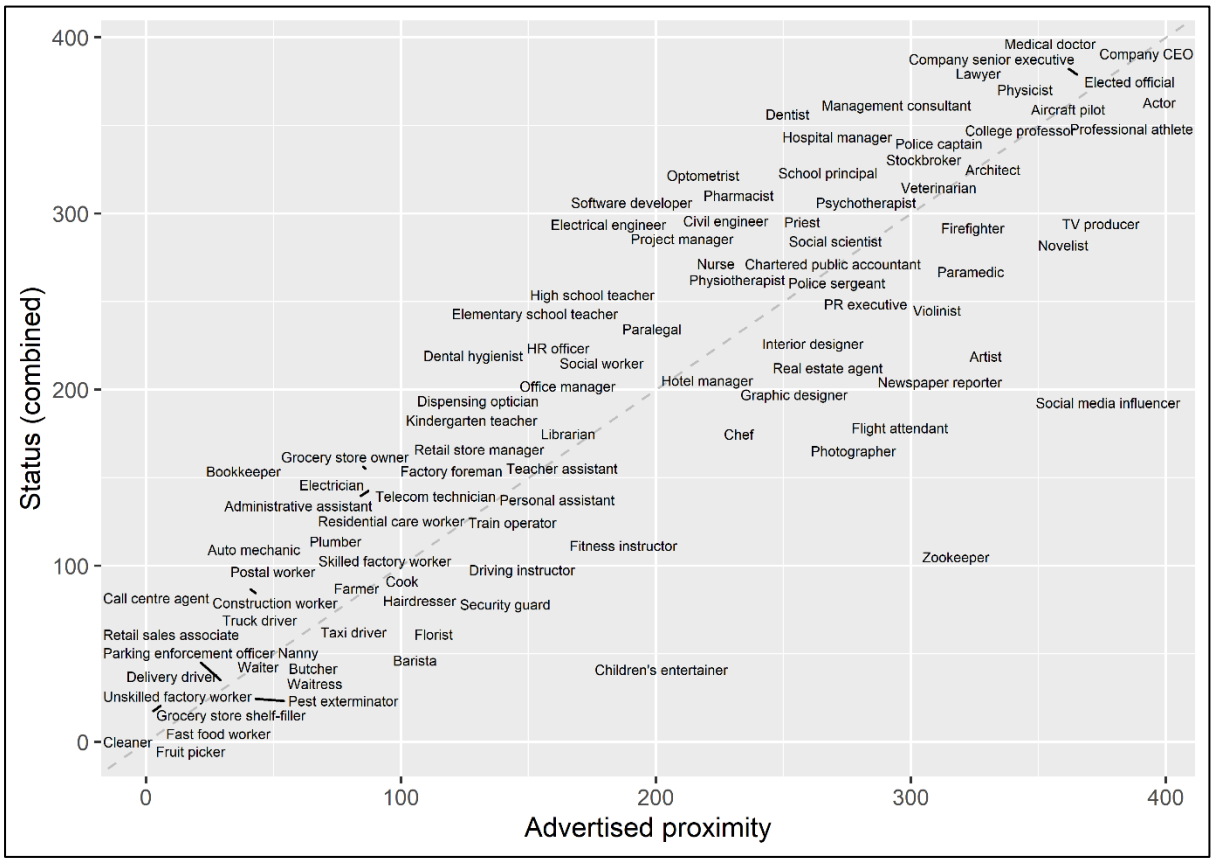

Figure D5. Relationship between the Status (combined) and Advertised proximity scales (GPT-4.1)

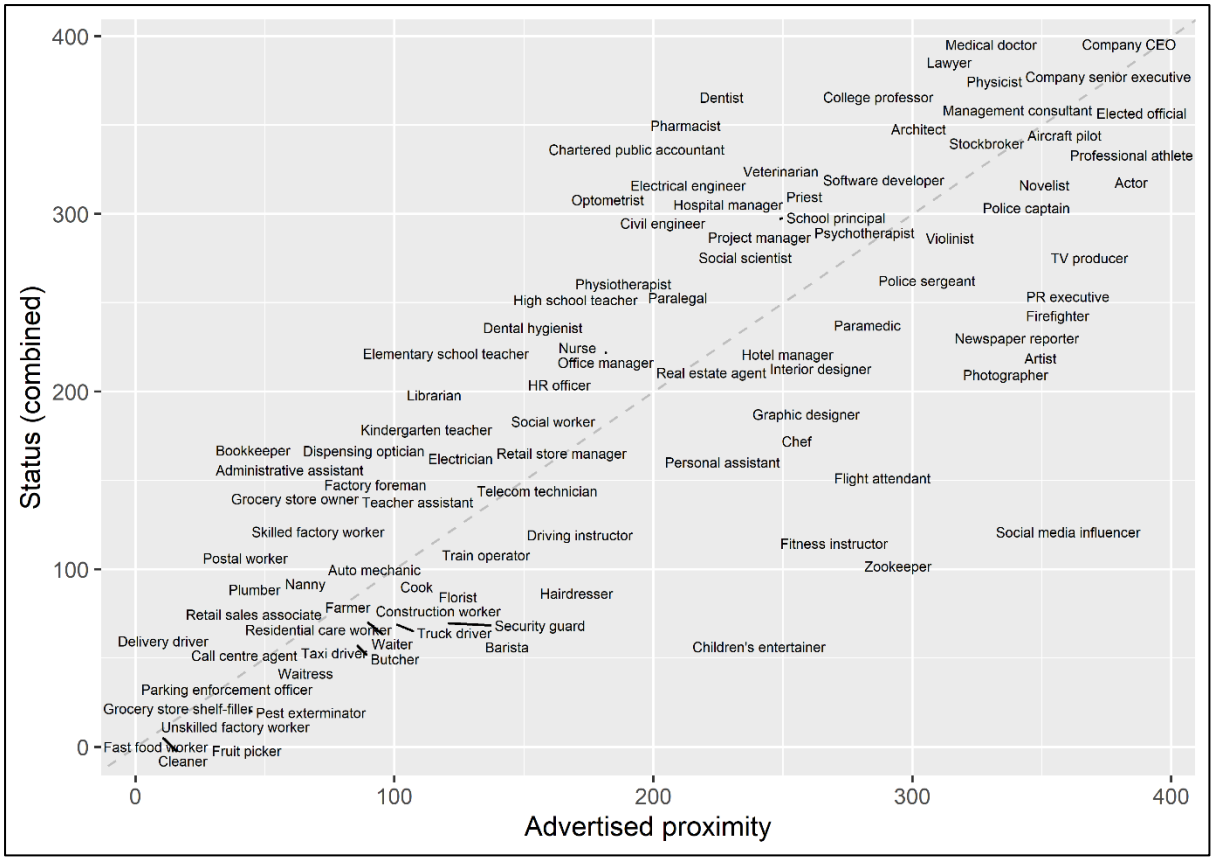

Figure D6. Relationship between the Status (combined) and Advertised proximity scales (Llama 3.3)

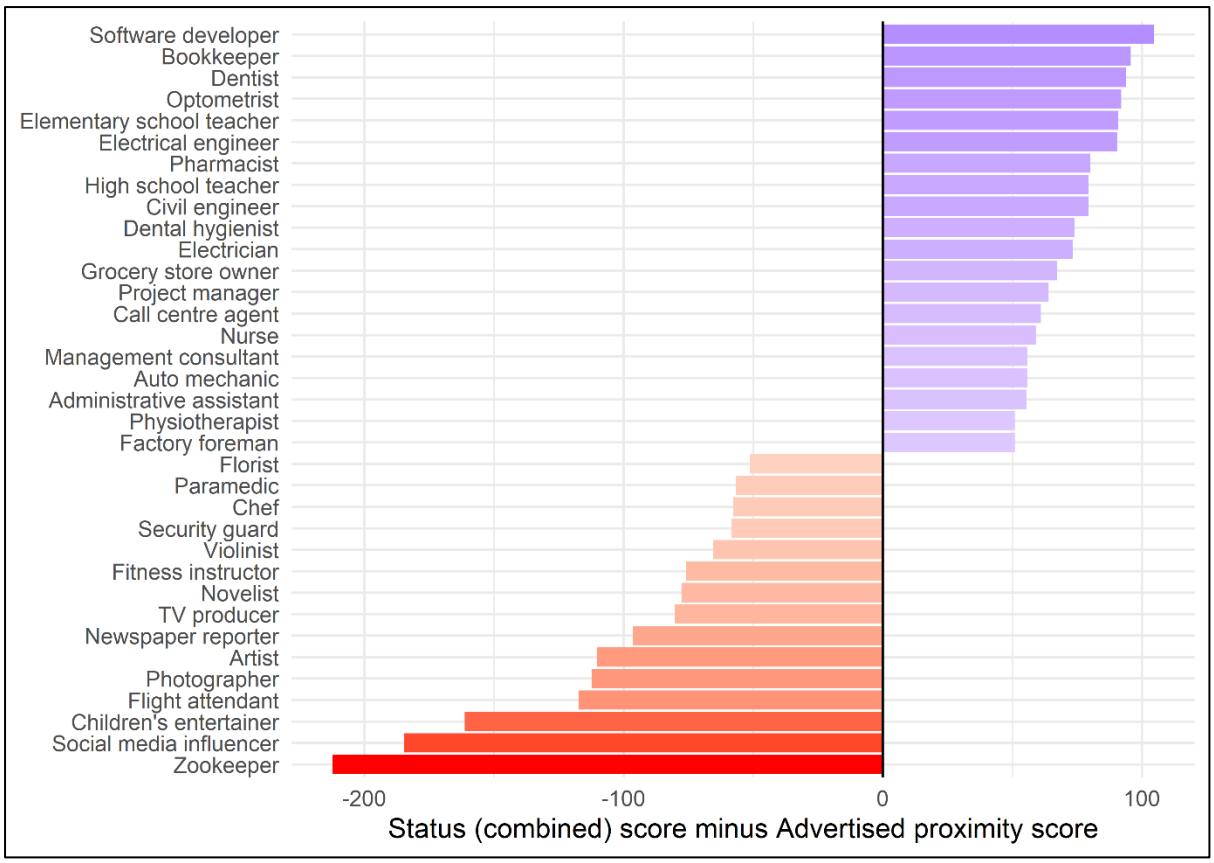

Figure D7. Occupations with the largest discrepancies between Status (combined) and Advertised proximity scores (GPT-4.1)

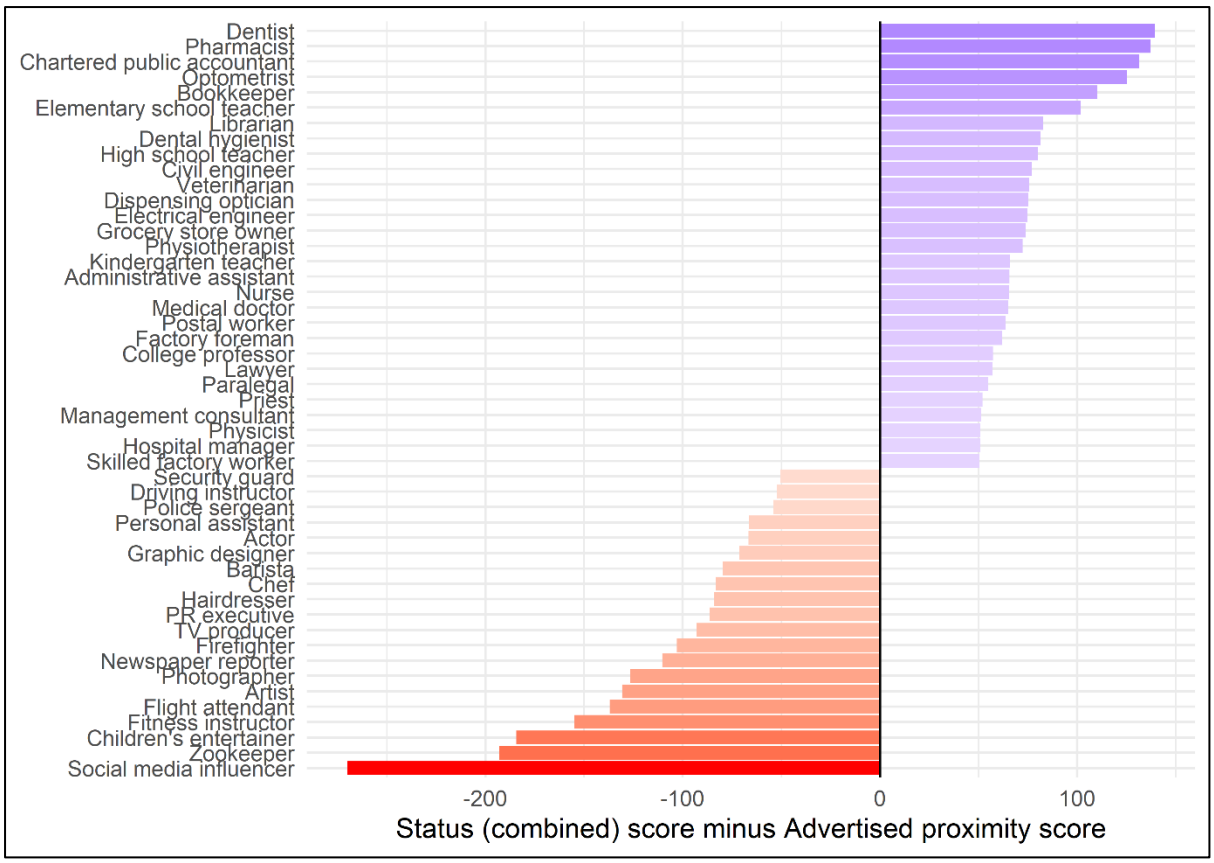

Figure D8. Occupations with the largest discrepancies between Status (combined) and Advertised proximity scores (Llama 3.3)

Appendix E: Comparison of ranks and scores for occupational characteristics (GPT-4o)

occupation	Training & pay		Social value		Authority		Unusualness		Public visibility		Excitement & interest	
	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank
Company CEO	392	1	121	67	378	3	328	15	368	5	285	29
Medical doctor	388	2	385	1	339	10	219	41	301	22	292	25
Company senior executive	382	3	78	79	371	5	246	33	306	21	241	37
Lawyer	379	4	175	51	367	7	99	68	323	16	252	33
Dentist	376	5	273	27	262	27	113	64	171	55	131	65
Management consultant	372	6	31	92	325	13	230	36	188	50	261	32
Physicist	368	7	302	19	260	28	357	7	212	47	301	22
Aircraft pilot	359	8	251	33	354	9	382	2	214	46	381	3
Optometrist	354	9	254	31	210	46	268	27	128	67	103	76
Stockbroker	349	10	5	98	231	37	272	24	188	50	286	28
Pharmacist	347	11	303	18	219	41	123	62	112	70	90	77
Hospital manager	342	12	293	20	360	8	242	34	238	36	167	55
Software developer	340	13	246	35	214	44	62	80	75	81	275	30
Chartered public accountant	336	14	128	65	271	23	210	43	105	72	57	85
Veterinarian	333	15	291	23	218	42	268	27	154	62	314	19
Electrical engineer	332	16	292	21	247	31	154	53	51	87	246	36
Civil engineer	323	17	328	12	313	14	191	46	170	56	227	42
College professor	321	18	313	15	297	19	226	38	281	28	232	38
Psychotherapist	318	19	304	17	234	34	333	14	119	68	232	38
Project manager	316	20	143	59	305	16	79	74	180	52	213	45
Architect	314	21	242	36	272	22	269	25	219	45	305	20
Professional athlete	312	22	34	89	121	67	391	1	384	3	387	2
School principal	308	23	347	10	330	12	149	54	323	16	128	66

PR executive	306	24	8	97	225	38	237	35	354	9	290	27
Physiotherapist	300	25	292	21	175	51	228	37	163	59	211	46
TV producer	293	26	45	86	301	17	350	11	345	13	359	9
Social scientist	290	27	276	26	250	30	325	16	244	35	264	31
Elected official	281	28	253	32	390	1	354	10	380	4	292	25
Police captain	279	29	314	14	381	2	278	23	347	11	340	12
Nurse	274	30	381	2	265	25	31	91	232	42	218	44
Dental hygienist	270	31	228	39	117	70	176	48	89	76	66	82
Police sergeant	268	32	312	16	373	4	168	51	320	18	328	16
Paralegal	264	33	153	56	234	34	205	44	85	77	152	60
HR officer	260	34	93	74	300	18	89	72	102	73	104	75
Real estate agent	246	35	33	90	204	47	67	77	327	14	251	34
Hotel manager	245	36	82	77	296	20	145	55	288	25	249	35
High school teacher	244	37	369	4	270	24	55	82	293	24	187	49
Interior designer	238	38	20	95	142	59	305	20	234	39	319	18
Office manager	237	39	62	82	243	32	43	85	80	78	55	86
Electrician	237	39	281	24	163	54	66	78	80	78	186	50
Graphic designer	236	41	58	83	85	76	189	47	168	57	321	17
Paramedic	231	42	367	5	338	11	224	40	275	30	356	10
Elementary school teacher	225	43	376	3	232	36	26	96	248	34	169	52
Telecom technician	221	44	207	47	124	66	198	45	36	91	158	59
Actor	212	45	20	95	114	71	357	7	392	1	391	1
Social media influencer	198	46	0	99	118	69	348	12	388	2	372	5
Librarian	194	47	215	42	140	61	264	29	78	80	73	81
Plumber	194	47	251	33	125	65	113	64	48	88	79	79
Retail store manager	193	49	115	69	214	44	28	94	224	44	114	69
Factory foreman	187	50	74	81	225	38	134	60	18	95	35	90

Social worker	186	51	363	7	264	26	135	59	234	39	230	40
Bookkeeper	185	52	40	88	147	58	121	63	12	96	14	96
Dispensing optician	184	53	190	50	106	72	320	18	72	82	53	87
Firefighter	182	54	365	6	370	6	254	30	294	23	378	4
Train operator	181	55	209	46	140	61	280	22	140	66	146	62
Grocery store owner	170	56	225	40	217	43	171	49	236	37	78	80
Newspaper reporter	168	57	263	30	121	67	310	19	349	10	341	11
Violinist	164	58	77	80	57	83	372	6	368	5	330	15
Auto mechanic	155	59	213	44	97	73	97	69	32	92	166	56
Administrative assistant	152	60	132	63	153	56	21	98	66	84	35	90
Truck driver	152	60	269	28	49	85	97	69	98	74	161	57
Personal assistant	137	62	28	93	132	63	104	66	107	71	161	57
Kindergarten teacher	134	63	356	8	165	53	97	69	158	60	169	52
Novelist	133	64	79	78	94	74	355	9	312	20	339	14
Photographer	133	64	48	84	86	75	253	31	317	19	360	8
Skilled factory worker	133	64	192	49	57	83	74	75	2	98	26	93
Chef	131	67	195	48	175	51	137	57	282	27	304	21
Flight attendant	131	67	149	57	259	29	252	32	346	12	362	7
Pest exterminator	128	69	161	54	77	77	325	16	27	94	114	69
Priest	120	70	214	43	313	14	378	4	284	26	113	71
Construction worker	119	71	267	29	149	57	35	90	113	69	168	54
Postal worker	113	72	239	37	72	79	100	67	142	65	42	88
Driving instructor	106	73	221	41	126	64	128	61	164	58	137	64
Fitness instructor	94	74	160	55	64	80	137	57	327	14	301	22
Call centre agent	92	75	32	91	41	90	37	89	48	88	33	92
Security guard	89	76	140	61	223	40	68	76	177	53	110	72
Zookeeper	88	77	147	58	195	49	382	2	145	63	340	12

Parking enforcement officer	88	77	28	93	278	21	285	21	262	33	16	95
Teacher assistant	82	79	327	13	179	50	56	81	155	61	106	74
Residential care worker	81	80	340	11	201	48	216	42	37	90	108	73
Hairdresser	74	81	103	73	61	81	31	91	231	43	219	43
Taxi driver	68	82	143	59	44	88	88	73	233	41	202	47
Cook	65	83	239	37	141	60	31	91	176	54	229	41
Delivery driver	63	84	171	52	10	97	27	95	198	49	120	67
Artist	58	85	106	71	77	77	346	13	367	7	368	6
Butcher	57	86	117	68	42	89	167	52	68	83	65	83
Retail sales associate	52	87	128	65	49	85	15	99	276	29	62	84
Farmer	45	88	350	9	242	33	170	50	57	86	117	68
Unskilled factory worker	34	89	106	71	16	95	51	84	1	99	2	99
Florist	32	90	45	86	40	91	269	25	143	64	178	51
Nanny	29	91	277	25	158	55	145	55	32	92	89	78
Barista	28	92	47	85	37	92	63	79	236	37	194	48
Waiter	26	93	90	75	61	81	43	85	273	31	150	61
Children's entertainer	24	94	132	63	12	96	376	5	355	8	296	24
Waitress	20	95	111	70	49	85	22	97	267	32	145	63
Grocery store shelf-filler	9	96	140	61	17	94	41	87	93	75	4	98
Cleaner	8	97	213	44	10	97	52	83	59	85	6	97
Fast food worker	6	98	85	76	31	93	41	87	204	48	24	94
Fruit picker	0	99	171	52	0	99	225	39	10	97	36	89

Appendix F: Comparison of Training and pay scale with Hauser and Warren (1993) Socio-Economic Index (SEI)

Figure F1, below, visualises the relationship between our GPT-4o-derived 'Training and pay' scale and the Hauser & Warren (1993) SEI scale. The SEI scale is functionally a weighted sum of 'occupational education' – the proportion of occupational incumbents in the 1990 US Census who had completed at least one year of post-secondary education – and 'occupational earnings' – the proportion of incumbents who earned more than a $25,000 a year (around $60,000 in 2024). Frederick (2010)[10] provides a table mapping Hauser and Warren (1997) SEI scores to IPUMS OCCSOC 2000 occupational codes. We therefore mapped our occupations to OCCSOC 2000 codes (and consequently to SEI scores) using the OCCSOC 2000 to OCC 2000 crosswalk provided by IPUMS.[11]

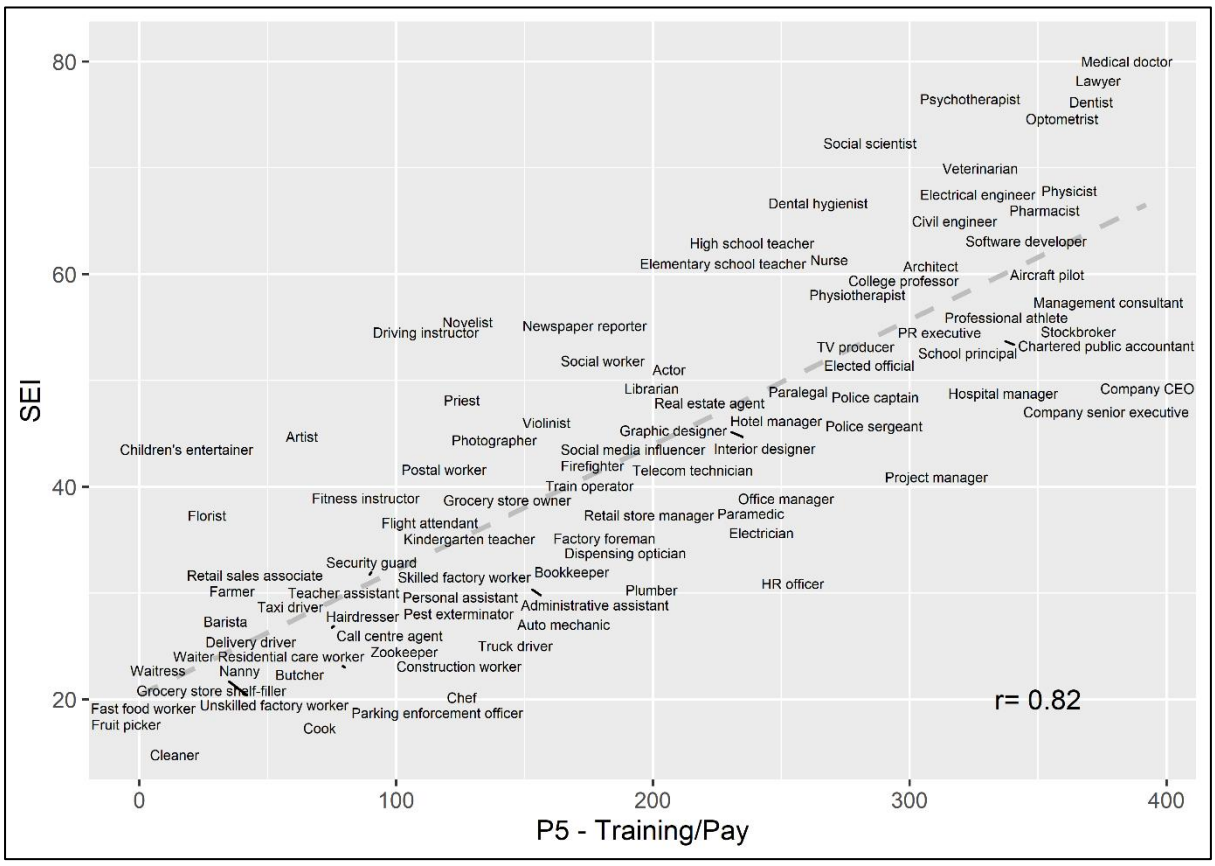

Figure F1. The relationship between our Training/pay scale (GPT-4o) and Hauser & Warren (1993) SEI

Figure F1 shows a number of occupations which GPT-4o rates as performing substantially better in terms of training and pay than would be suggested by their position on the SEI scale. These include, a number of blue collar roles, such as *Chef*, *Auto mechanic*, *Electrician*, *Plumber*, and *Truck driver*. This may be a consequence of popular narratives around the relatively high pay

[10] Frederick, C. (2010). A crosswalk for using pre-2000 occupational status and prestige codes with post-2000 occupation codes. Center for Demography and Ecology: University of Wisconsin-Madison. PMID, 25506974.

[11] https://usa.ipums.org/usa/volii/occtooccsoc18.shtml

associated with these roles. Similarly, our training and pay scale appears to under-estimate the pay/training of, notably: *Novelists*, *Artists*, *Newspaper reporters*, *Social scientists*, and *Teachers*. These discrepancies may arise due to how 'training' is interpreted by GPT-4o in the construction of this scale. Specifically, the prompt underlying our training and pay scale requests the education/training *required* for an occupation; whereas SEI scores are based on actual qualifications held. This may explain why artists, novelists, and newspaper reporters perform so much better on the SEI scale. These occupations do not strictly *require* formal qualifications, but are nevertheless dominated by college graduates.

Appendix G: Comparison of ‘Status (combined)’ and ‘Training and pay’ scales

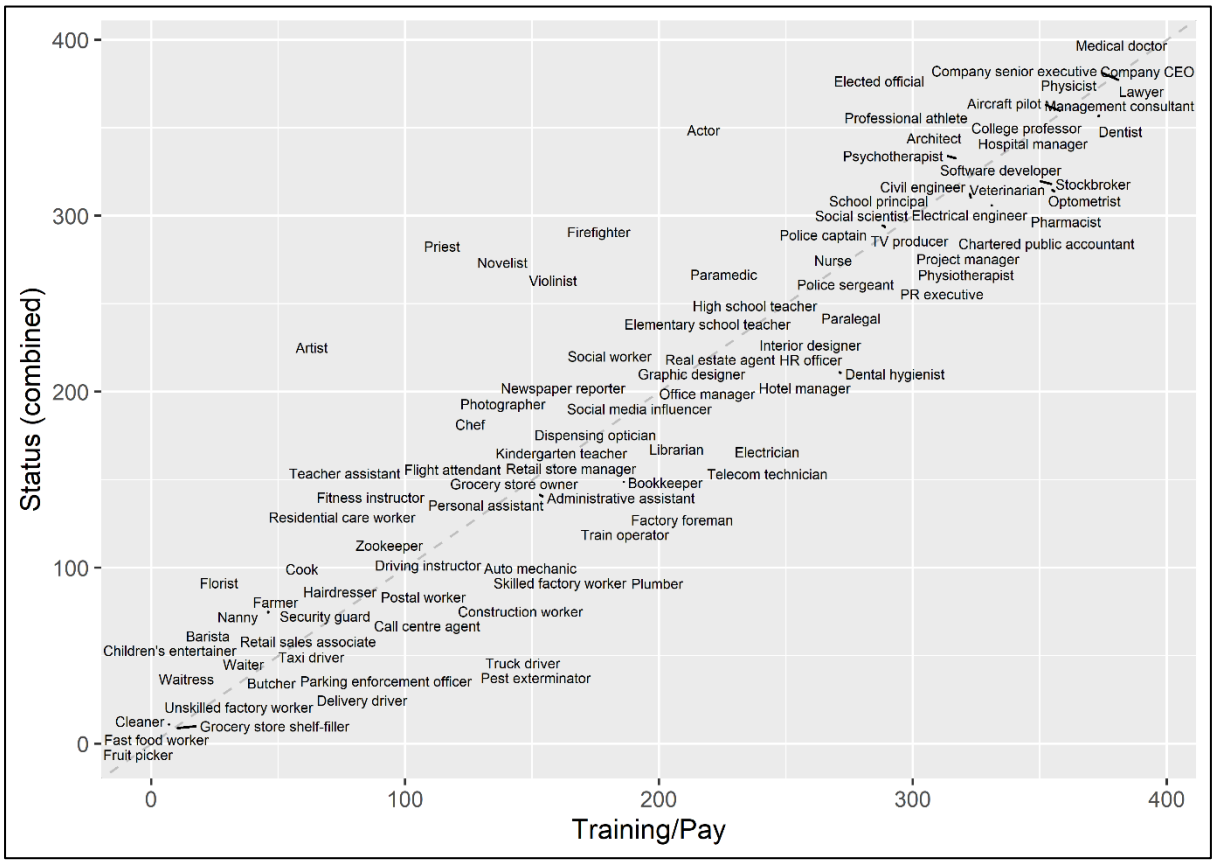

Figure G6. Relationship between our ‘Status (combined)’ and ‘Training and pay’ scales (GPT-4o)