Evaluations

Funcwords. This provides initial evidence for the integrity of the semantic relations captured by the word embeddings, and that these semantic relations can be compared across word embeddings corresponding to time periods.

Analogy

Some of this is due to their distance in time to the analogy formulations - language has changed substantially since 1900. But the semantic relations asked of in analogy tasks are relatively stable over time - London is to England as Paris is to France and will likely remain constant for a very long period of time. This is also clear in the results - the difference in performance from 1900 to 1990 are minor - 0.13 in ENG\\_ALL and 0.08 in COHA, and no increase in ENG\\_GB or ENG\_FIC. Performance, instead, seems to correlate with sheer size of the corpus, with the rank in analogy performance matching the rank in corpus size. This would also explain why the state-of-the-art embeddings are achieving higher performance - they are much larger than the corpora used in this study.

This comparative lack of performance may bear consequences, since the ability to solve analogy tasks is what the logic of word embeddings semantic analysis is built upon; geometric relations mirror semantic relations. Granted, analogical reasoning is only a subsection of all semantic relations, but it is an important proxy.

Word pairs

The ranking in performance points to size being the most important determiner of ability to match human rated word senses. Interestingly, however, time is a much more important factor here, with all embeddings but

This implies that intuitions of word similarity change much more than analogical relations, which perhaps remain more constant because of their more fundamental relations of meaning.

Gen

This is positive for our study for two reasons. First, the relatively good performance of these word embeddings implies that they can capture these more subtle relations of meaning, which is exactly what is hoped for with an analysis of class connotation. Secondly, their increase in performance implies that subtle relations of meaning change a substantial amount over time, indicating that word embeddings have something to model. The latter conclusion may have to be tempered, since the drop of the performance may be due to the relative smallness of the early decade corpora. But considering the relatively large increase in performance compared to the analogy over the same decades, as well as the relatively small difference in size of the 1900 to 1990 corpora, the conclusion still seems to hold.

Dimensions stability

The large fall in correlation for education could be attributed to the changing status of education in the 20th century. At the beginning, many fewer went through a protracted period of education - university being restricted to the very few. But as access to education increased, as well as university educations becoming a necessary step to a significant portion of jobs (i.e. the professions), the role education has played in relation to the world of work has transformed. Consequently, one can imagine that it has attained a different connotation over time.

Employment, likewise, was something that gradually became necessary for the upper class over the 20th century. At the inception of the century, most of the upper class could and would avoid employment, living off of their accumulated capital and proceeds from the capital such as rent. But gradually, it became necessary, as well as the norm, to be employed. Hence the list of the world's richest is now populated by working men (who also commonly preach the values of hard work).

Gender aligns with expectations from the historical record. While it remains relatively stable in the initial decades, it begins to drop precipitously in correlational value in the 1960's. This aligns directly with when the feminist movement picked up momentum and transformed cultural expectations of women.

That status remained constant is surprising. With the cult of celebrity that came with popular media in the latter decades, one would expect that high status would have different connotations. It may also be that the difference between words included in that antonym list like "famous" and "respected" shows that difference, but this gets tempered in the procedure of aggregation. It is also surprising that cultivation remains quite stable - theories of omnivorism have suggested that distinction according to cultivation is becoming more subtle in recent decades. One could imagine, however, that words like "cultivation" are not being used to describe the new way of having good taste - words like "hip" or "cool" or "fresh" displacing that and cultivation remaining ossified and tied to an old ideal. That affluence remains similar is perhaps more understandable - although one might expect that the increasing prevalence of the concept of the "new rich" in contradistinction to old aristocratic wealth might change that.

Dimension validity

The survey correlations show a competence in the cultural dimensions capturing connotations of words. Gender seems to remain a benchmark, while affluence and race trail behind. As discussed, race can be discounted because of its polysemy. Affluence, however, may have fallen due to a lesser prevalence of class markers by affluence - meaning a greater variability amongst individuals' associations with wealth (and thus declining ability of embeddings of match associations). This is corroborated by the average variance for the respondent ratings for affluence which, at 314.438, is higher than race (313.264) and gender (284.038). Gender is surprising, however, as one would expect a fall in correlation with gender associations, consider how feminism has opened up more variability in women's lives and thus greater variability in bias and representation. This may point to enduring biases in that domain.

\*\*Caveats to further analysis\*\*

The aforementioned analyses have shown that the embeddings have an unevenly distributed reliability. ENG\\_ALL is the most reliable of them on all accounts, followed by ENG\\_GB, ENG\\_FIC and then COHA. They seem to perform reasonably on both analogy and word pair similarity tests, but are far from perfectly representative of human semantic relations. Affluence and gender connotations seem reliably captured, but all considerations of race will have to be discounted because of the inconsistency of the correlation evaluation results. Their reliability is harder to verify the further back in time we go, as there is a lack in historical surveys to test them on. We can assume, however, that a significant amount is captured – both because of the relative performance of the earlier word embeddings, but also because of the validation that Kozlowski et al. performed with the 1950’s semantic differential research.

# Lifestyle homology

(of gb12s poor performance) The best explanation seems to be the relative size of the corpus trained on, as a similar drop in accuracy was noticed in the previous rounds of analysis. Because of the relatively poorness in performance of embeddings in the initial evaluations, it was decided that they were not robust enough for the fine quantitative analysis that the subsequent analyses required.

Poor performance gss) This may be because of the polysemy of many of the words (does "art" refer to making, visiting, or discussing art? Does "sport" refer to playing sport or watching sport? Could connotations be more relevant to artists or sportsmen than their consumers?). It may also be due to the lack of specificity of the activities. As cultural capital researchers ¤have argued, too broad a category like "sport" is not enough to map distribution of taste - type of sport, or even choice of sports team are more relevant to differentiation on taste.

GBCS performance iactivites) This is confusing, since there doesn't seem to be a great degree of recognisable difference between the two collections of words. It may be because of the Great British Class Survey being conducted in the UK, where actual class practices are more distinct than in America, where there is more fluidity in practices and less distinct boundaries for each class. ¤BUT WHAT ABOUT ENG\\_GB??

Music & gss: This implies that connotations of affluence with music in the word embeddings reflect actual connections of music taste with affluence.

Music &GBcs: These must be taken with a certain degree of caution, since only 4 words were retrievable from the word embeddings after filtering out the very polysemous genre words (e.g. country, rock and metal)

GENERALLY:

The fact that word embeddings can replicate real world associations between affluence and music taste is interesting. This would indicate that music is still a type of cultural consumption that is inherently classed. Collectively, there seems to be an awareness which music genres are engaged listened to by the rich and which by the poor. This aligns with Bennet's (¤culture, class distinction) findings in his survey on cultural capital and social exclusion, where music found to be the type of cultural consumption which differentiated people most (above book genres, film genres and art genres).

There seems less of an association with cultivation and status, however. Cultivation perhaps, is less relevant a means of differentiating people's music taste - this would align with the omnivorousness hypothesis of the upper and middle classes being less confined to high culture and instead listening widely. It may also be that cultivation and associated terms are used less to describe music taste, with their connotations of the old, universal humanist notions of Bildung.

Status seems inherently problematic because of the words used. One can imagine that words like "respected" and "famous" are much more likely to refer to the artists that produce works in these genres, than the types of people that listen to them.

The correlational analysis is not without its flaws, however. Income is not a perfect proxy of one's socioeconomic status. Especially when moving up the higher ranks of socioeconomic status, wealth is increasingly determined by one's accumulated capital and less by income. Because of this, the correlations face a potential mismatch of scale - the upper end of the survey values referring to the wage earning upper middle class and the upper end of projections referring to either the upper middle classes or the traditional upper classes (each holding a distinct position in the cultural space). Likewise, the measure of education is not linear, as the ranking code implies. Perhaps years of education would have been a better way of linearising, or a simple university versus non-university binary.

Furthermore, taking the mean distracts from the variance there may be in each sample. If the variance is high, it would be hard to assert there is a good mapping, since only few values of the sample would genuinely accord with the values.

¤CHECK VARIANCE OF EACH MEAN VALUE CALCULATED

- Evals of both: polysemy, participation vs taste, the problem of averaging and not considering variance

# Space of lifestyles

Now we take the list of leisure activities present in the GBCS and GSS surveys and combine it with three further lists - places of entertainment, sport and clothing (music was discarded for having too many recently established genres), and analyse how these domains transformed in their class connotations through time. We use the same correlation-of-projections method used earlier in the validation of cultural dimensions on each of the domains, correlating the projections of the 1900 word embeddings and the projections of the 1990 word embeddings.

We can interpret high statistically significant correlations as measuring the consistency of a domain's space through time. A high, statistically significant correlation implies, then, that the relations between the objects of the domain remain stable with respect to one another, in terms of each of the cultural dimensions.

Inspecting their graphs (see appendix), it does not seem to be because of a dramatic change in the space - the correlations are very low from year to year to begin with - which implies that i) they are not relevant in the consideration of sports, or ii) they are not particularly meaningful dimensions. Considering the relative performance of each dimension in the dimension stability tests, it seems to be that the first applies for education, and the latter applies for race.

Sport seems to be remarkably consistent throughout the 20th century. This implies that the class connotations of each object seem to not change, which is surprising considering the opening up of many sports to the working classes because of increasing disposable income. Clothing also seems remarkably consistent - which one would not expect because of the decreasing distinctiveness of clothes worn by each class over time (tokido). Though its lowest correlation is in the dimension of affluence (0.463), implying that there has been some mixing up of the space. The domain of leisure seems to have transformed widely - except for gender and affluence, which seems to accord with expectations, since many activities of leisure are prohibitively expensive (opera, restaurants) or tastes that are more-or-less exclusive to the rich (classical music, theater). Places of entertainment seem to display a great divergence over the century, which again may be due to the increased disposable income available to most people, and the consequent intermingling of people at places of entertainment.

Comparing these correlations with those of the top 10,000 most frequent words seems further confirm this picture. The top 10,000 most frequent words show a statistically significant correlation in all dimensions (albeit very low in race and education, as seen in the other domains). This indicates that there is a general stability, from which the domains can diverge – the insignificant correlations indicating, then, a divergence from this general stability.

It is hard to justify interpreting too much, from this mode of analysis, however. The argument I have been making - instability in a particular domain compared with a greater stability generally implying a transformation of that domain’s space - may have other explanations. Small sample of words from the domain may mean a higher risk of insignificant correlations. Likewise, speculating reasons for these changes is also risky, however, as multiple narratives could fit the same result. Deeper engagement from a domain expert would greatly contribute in providing contextual depth.

# Space of lifestyles

. It is the most popular of the art forms, and thus most prevalent across classes.

The theory of working classes greater engagement with the physical, combative sports ¤bourdieu quote is badly reproduced – wrestling scores comparatively low on both cultivation and affluence, but boxing scores high on the affluence dimension (though this may also be due to the comparative remuneration of the celebrity boxers of the 1990's). Football, the most popular sport for the working classes of Britain (though, of course, polysemous with American Football), scores very high on affluence and cultivation dimensions. On the whole, sports do not seem to be particularly well captured by the projections - but unfortunately no datasets were available to test their performance on.

In sum, it seems that the space of music genres is best captured by the word embeddings, followed by leisure and then sports, and finally places of entertainment. What is evident, however, is that in many of the domains, they seem to be right in \*most cases\*, as evidenced by their relative mean positions, but are not particularly useful for analysis of individual objects. If we were to learn about the elative position of each of the lifestyle domains only from these graphs, we would have a quite warped understanding of society.

# Occupations

Analysing the space of occupations had a similar structure of approach to analysing the space of lifestyles. I will begin by comparing the projections of the occupation space on the 1990 data from the US census with a correlational analysis. I will then consider how to interpret these correlations, with special attention given to outliers. Following this, I will examine this for previous decades. I will then proceed to visualise the space through projections onto cultural dimensions. I will compare these visualisations with visualisations of projections from principal component analysis. I will then consider how different occupational categorisation schemas perform in capturing popular conception using silhouette analysis. I will finally examine the individual cases of the manager and professions, as examples of more in depth study using this methodology.

It may also be the effect of education having a more practical function as training for employment in the late 20th century - rather than the humanistic education that was the privilege of the elites before.

# Outliers

Care was taken to remove outliers from the dataset - as Pearson's correlation is extremely sensitive to outliers and thus assumes a dataset without them. ¤REPOSITION AS AN EXAMPLE AREA TO GO INTO

Prior to removal of outliers, affluence was not statistically significant, but after manually removing those that diverged greatly from the general trend, it became significant. These were predominantly those in the artistic domain - dancers, designers, painters, authors, athletes and photographers. These were all projected much higher on affluence than were actually compensated. One can imagine that these occupations are naturally more associated with affluence because of the celebrity of their most prominent representatives. Likewise, dentists, physicians, surgeons, policemen, lawyers, judges, pilots, veterinarian and stock and bond salesmen were removed. These were all represented as much poorer than they actually were recompensed. It is unclear as to why that may be the case, but for some like dentists, pilots and veterinarians, their compensation is arguably much less represented by the media and consequently less associated with wealth. Others like lawyers, judges and policemen often are associated with criminals because of the nature of their work. Criminals tending to be poorer, it would seem reasonable to assume that these words often neighbour words describing criminals, and therefore associated.

For education, there were no prominent visible outliers, so it was left as it was found.

Status had a few prominent outliers, however. Members of the armed forces and painters were much higher projected than were ranked as prestigious. Tool makers, die makers and setters were projected substantially lower, however. This divergence cannot be attributed to difference in public perception like the previous because the rankings are themselves subjective. It may be attributed to difference in the perception of those writing in the public domain to that to the public perception, but that also has to be discounted, as occupational prestige has been demonstrated, as mentioned previously, to be consistent across demographic subgroups. It would most likely be due to statistical noise specific to these words - but further, qualitative research or a domain expert's opinion would be needed to judge whether this is the case.

Whether outliers should be removed in this case is debatable. These could arguably be due to deficiency in the measuring instrument (in this case the shortsightedness of the algorithm). They could also be capturing actual popular perceptions of occupations. It may be the case that deviations represent actual deviations of popular conceptions from reality. It may well be the case that the common person considers artists to be rich, with the majority of media exposure being of the well known and accomplished and less of the starving artist trope. These speculations can only go so far, and it would require further inspection by a domain expert - perhaps a cultural theorist or historian - that could assess reasons for divergence. This points to how word embeddings could be used as an exploratory tool, whose outliers can prompt investigation into prominent stereotypes.

General

That there was a high correlation between the embeddings and status is very promising. Unlike the proxies for education, cultivation and affluence, the values for status were aggregates derived from subjective rankings by occupational prestige research. This suggests that not only are word embeddings competent at capturing objective relations or properties of the world, they are even better capturing subjective impressions of cultural objects than objective properties or relations of the world. This discrepancy is important, as it implies that word embeddings can, in fact, be used to evidence stereotypes or biases, considered as unrealistic representations of the world.

# Space of occupations

With the projections partially validated, I will now proceed to examine how the projections of the occupations distribute across the different class dimensions. As

¤, I assigned each occupation to a class or category, following the International Standard Classification of Occuaptions, and coloured them accordingly. The words are far from one-to-one, with some being belonging equally to different groups, and many being left out, due to the restriction of being able to use only one word. The visualisations

(on artists) As aforementioned, some of this may be due to the bias of celebrity that these domains tend to accumulate to themselves. If we consider them from the perspective of perceived social position, they seem to occupy a distinct position in the social space. This accords more with Bourdieu's multidimensional conception of the distributions of occupations of social space. In contrast to the traditional reliance on economic capital, the Bourdieusian system asserts cultural capital and social capital (as imperfect a proxy "status" is for social capital) also to be important differentiators of the space of occupations.

# Professions

To investigate the domain of professions, we first consider how well this domain correlates with the census data. Correlating projections of each occupation on the 3 dimensions of affluence, status and education with the mean average of income, occupational prestige and level of education, we get surprisingly unconvincing results.

This is hard to interpret the meaning of. At the very least, it implies that word embeddings are relatively unstable when considering smaller subsections of a whole.

We turn to consider how the image of professions has changed through time. We create for each decade a composite vector by calculating the mean average of all of the vectors corresponding to the professions. We verify that this composite is meaningful by looking at the most similar vectors (by virtue of cosine similarity) and seeing whether they were related to the professions. We then project each decade’s composite vector onto each of the cultural dimensions, and plot each of the dimension projections through time as separate lines.

We now look at the space of professions by projecting them onto the cultural dimensions of the 1990 word embeddings

– perhaps because their more practical nature meaning they are less associated with high education and status.

This is an inherent issue with the bias of landscape graphs weighting towards horizontal consideration.

This indicates that affluence seems to be the axis upon which professions differentiate – and it seems to be that this is due to the type of work that is involved – intellectual, practical or artistic. This is, of course, a relatively crude method of evaluating this space and further research would be required with the knowledge of a domain expert to parse the data in a more meticulous manner.

# Manager

We now take to task the New Spirit of Capitalism's assertion of the transformation of the spirit of capitalism by analysing the image of the manager in word embeddings. We hypothesise that the word "manager" and its synonyms become greater associated with education, because of the institutionalisation of business management in the universities. We hypothesise also that it becomes more associated with affluence and status, as managers become increasingly well remunerated (especially with respect to the role of CEO) and are accordingly provided status. Likewise, we hypothesise an increasing association of the word with the word space described for the new spirit of capitalism (connectionist words like network, projects, team, flexibility, lean, and so on - full list provided in the appendix ¤).

This implies that the word “manager” captures the changes in the profession over time. The other words in the composite vector – foreman, supervisor, superintendent – these became perhaps more obsolete over time. Being displaced by manager, they likely did not change as much in connotation over the period concerned. For the ensuing analyses, we therefore restrict our focus to the word “manager”.

**Subcats**

* Micro level: spaces, professions & managers thru time