

Are the Beautiful Good in Western Literature?: A Simple Illustration of the *Necessity* of Literary Quantification

**Jonathan Gottschall, Christine Callanan,
Nicole Casamento, Natalie Gladd,
Kristen Manganini, Tanya Milan-Robertson,
Patrick O'Connell, Kimberly Parker, Nathan Riley,
Valerie Stucker, Adam Tapply, Christopher Wall,
Alexis Webb**

Summary

Social psychologists consistently find evidence for a physical attractiveness stereotype (PAS) whereby people attribute positive personality characteristics to the physically attractive and negative personality characteristics to the physically unattractive. Researchers summarise the PAS phenomenon with the aphorisms “what is beautiful is good ‘and’ what is ugly is bad”. Are traces of this stereotype detectable in Western literature? Are “good” characters more likely to be portrayed as physically attractive and are “bad” characters more likely to be portrayed as physically unattractive or ugly? And what is the best way of addressing these questions? This study presents results of a quantitative content analysis designed to determine whether or not the PAS is reflected in a sample of plot and character summaries taken to be reasonably representative of “canonical” Western literature. Results are reported for the sample as a whole, based on the characters’ sex, and for three broad historical periods. Discussion focuses on the relevance of our results to debates about the origins of the PAS, and on the unexpected finding of sharp differences in the strength of the stereotype in male versus female characters. More broadly, we seek to provide a modest illustration of the value of a more scientific approach to literary study.

Opsomming

Sosiale sielkundiges vind voortdurend bewyse vir 'n stereotipe wat berus op fisiese aantreklikheid (physical attractiveness stereotype, of PAS), waarvolgens mense positiewe persoonlikheidstrekke aan fisies aantrekklike mense toedig, en negatiewe persoonlikheidstrekke aan fisies onaantrekklike mense. Navorsers som die PAS-verskynsel op met die aforismes “wat mooi is, is goed” en “wat lelik is, is sleg”. Is daar spore van hierdie stereotipe in Westerse literatuur te bespeur? Is dit waarskynlik dat “goeie” karakters as fisies aantrekklik en “slegte” karakters as fisies onaantrekklik of lelik uitgebeeld sal word? En wat is die beste manier om hierdie vrae te benader? Hierdie studie lê resultate voor van 'n kwantitatiewe inhoudanalise wat

ontwerp is om te bepaal of die PAS weerspieël word deur 'n monster van plot- en karakter-opsommings wat beskou word as redelik verteenwoordigend van “kanonieke” Westerse literatuur. Resultate word vir die monster in sy geheel weergegee, gegrond op die karakters se geslag, en vir drie breë historiese tydperke. Bespreking fokus op die relevansie van ons resultate vir debatte oor die ontstaansmoontlikhede van die PAS, en op die onverwagte ontdekking van skerp verskille in die sterkte van die stereotipe in manlike versus vroulike karakters. Op 'n breër vlak streef ons na 'n beskeie illustrasie van die waarde van 'n meer wetenskaplike benadering van literêre studie.

Introduction

Social psychologists consistently find evidence for a physical attractiveness stereotype (PAS) whereby people attribute positive personality characteristics to the physically attractive and negative personality characteristics to the physically unattractive. Research subjects consistently prejudge attractive people as more likeable, happy, successful, sexually warm and responsive, kind, socially graceful, occupationally competent, and intelligent than less attractive people. The stereotype is already well established in the minds of schoolchildren, and perhaps even in infants, judging from the length of time they stare at attractive versus unattractive faces. As a result of this “lookism”, people judged to be physically attractive by panels of independent raters receive better grades, land more prestigious and lucrative jobs, win more arguments, are more popular, marry wealthier spouses, receive lighter prison sentences for the same crimes, make the most successful salespersons and, not surprisingly, report being happier (for overviews, see major meta-analyses by Feingold 1992; Eagly, Ashmore & Makhijani 1991; Langlois et al. 2000). Thornhill’s review of the PAS literature leads him to opine, “There is probably more prejudice in the context of physical attractiveness per se than on the basis of ethnic group and sex combined” (1998: 559). In short, evidence in favour of widespread and mostly unconscious prejudices in favour of the physically attractive and against the physically unattractive has been accumulating for decades, and is now so well established as to approach the level of social science verity.

On the surface, there appears to be good evidence that traces of a PAS are detectable in literature. Literary characters who are both beautiful and good are legion. A few memorably unattractive “bad” characters include Shakespeare’s Richard III, Homer’s Thersites, Dickens’s Uriah Heep, and Cinderella’s stepsisters. Moreover, the hideous appearances of monsters, ghouls, goblins, gargoyles, and witches symbolise their unwholesome natures. On the other hand, examples may almost as readily come to mind of characters who violate this pattern. The physical appearances of Quasimodo, Cyrano de Bergerac, and numerous other unattractive

characters belie their essential goodness, while the beautiful Queen in “Snow White” and lovely Duessa in the *Faerie Queene* are but two in a long line of femme fatales who approach evil incarnate. In short, in the Western literary tradition the beautiful are obviously not always good, and the ugly are not always bad.

So the task of determining whether there *is* a PAS in literature – not in the sense of being represented in individual works but as a salient phenomenon running through whole literatures and genres – may not be as straightforward as it first appears. When queried whether the stereotype is broadly reflected in literature, the scholar must respond by cudgelling his or her memory for examples and counter-examples and, in that fashion, try to provide an accurate response. But, when awash in the whole universe of literary characters, amidst all the characters who fit the stereotype and violate it, how is it possible to arrive at a firm answer? Any scholar citing a proliferation of characters who fit the pattern can be countered by another scholar citing droves who do not. The winner of this hypothetical debate will not necessarily be the scholar who more closely approximates “the truth”, but will just as likely be the scholar who is more rhetorically fluid or who happens to be able to conjure from memory a larger or more resonant set of examples.

In a debate of this kind the traditional qualitative methods of the humanities collapse beneath the burden of selection and confirmation biases. Believers and non-believers in the reality of a literary PAS can volley examples and counter-examples, but it is difficult for them to actually determine if the phenomenon exists in literature or if it is only a phantom produced by an unrepresentative set of examples.

It may be objected that we are pushing too hard. Perhaps many literary scholars would accept that something like a PAS operates in literary works, and that the many counter-examples represent cases of authors toying with the prejudices of audiences. But the implications of our argument radiate beyond the confines of attractiveness prejudice. The problems of confirmation and selection bias discussed above are not unique to the question at hand; they are recognised as substantial obstacles to the progress of knowledge in literary analysis.

Our most immediate goal in this article is to determine, using methods of quantitative content analysis developed by social scientists (Neuendorf 2002; Krippendorff 1980; Weber 1990), whether or not the PAS exists in a literature sample taken to be reasonably representative of “canonical” Western literature. On the whole, are “good” characters more likely to be portrayed as physically attractive and are “bad” characters more likely to be portrayed as physically unattractive? But we also have larger goals. This article uses a narrow and relatively simple question about attractiveness

prejudice to explore the potential applicability of a methodologically scientific approach to certain broad classes of literary questions. We seek to provide a simple and direct illustration of why – for the discipline of literary studies – a *minimally* adequate methodological toolkit requires more substantial quantitative components.

We are far from the first to make an argument for quantitative methods in literary study or to attempt to demonstrate their power (e.g. Schmidt 1982, 1992; Martindale 1990, 1996; Bortolussi & Dixon 2003; Gottschall 2005; Moretti 2005). Learned societies like the International Association of Empirical Study of Literature and The International Association for Empirical Aesthetics, as well as journals like *Poetics*, *Computers in the Humanities*, and *Empirical Studies of the Arts* have long promoted quantitative methods. However, these journals and societies mainly serve social scientists who are interested in aesthetics; quantitative work still represents a tiny fraction of the total output of literary scholarship, and what there is of it languishes far outside of the mainstream.

However, the point of this article is not to attack traditional qualitative methods of literary study. Nor is the point to sing paeans to quantitative methods. Literary quantification has salient limitations and pitfalls; it can only serve to complement traditional approaches, never to replace them. In the end, the question is not of qualitative or quantitative methods being inherently superior, but of applying the right methodological tools to the right types of problems. To argue for the superiority of quantitative over qualitative approaches (or vice versa) would be as vacuous as arguing that hammers are better than drills. As the carpenter requires a collection of widely varied and subtle tools for effectively confronting widely varied challenges, so too does the scholar. The question of whether or not there is a PAS operating in a given work or small group of works might best be approached qualitatively. Quantification might prove clumsy, crude, and unnecessarily complicated. However, if the question is extended to encompass large sets of works or whole traditions (e.g. the Western canon) then qualitative approaches can quickly be overwhelmed in floods of data. These large-scale challenges represent one area where quantitative methodologies, designed to confront large data sets and to control for selection and subjectivity biases (while never fully eradicating them), may often be most appropriate.

Data and Methods

We undertook a content analysis of 240 complementary summaries of plot and character information from works of “classic” Western literature in

order to test the following prediction generated from theory and research on the PAS (e.g. Dion et al. 1972; Eagly et al. 1991): *protagonists will be both more likely to be defined as physically attractive, and less likely to be defined as physically unattractive, than will be antagonists*. This requires some explanation. For the purposes of this research we defined a protagonist as a character who plays a central role in the action and for whom the audience is led to root predominantly for rather than predominantly against. An antagonist was defined as a character who plays a central role in the action, who acts as an obstacle to the goals of the protagonist(s), and for whom the audience is led to root predominantly against rather than predominantly for. The clear assumption inherent in this prediction is that protagonists will be, as a general rule, far more likely to be portrayed as “good” than antagonists, despite some exceptions. In other words, research on the PAS suggests that characters’ physical attractiveness will correlate with their agonistic status: a pattern should emerge whereby characters with more positive behaviours and personality traits (protagonists) are depicted as more physically attractive than those with less positive behaviours and personality traits (antagonists).

Coders were instructed to provide attractiveness information only on the basis of explicit judgements in the texts; they were told to avoid providing information based on their individual attractiveness preferences. In practice, this means that codings were overwhelmingly based on explicit applications of synonyms for physical attractiveness or unattractiveness to given characters (e.g. beautiful, handsome, ugly, repulsive). However, coder judgement played a substantial role in many cases. If, for instance, a character was described as “repulsive” the coder would have to judge, based on context, whether the character was being described as physically ugly or as repulsive in some other sense (e.g. morally repulsive). If the context did not indicate repulsiveness in the relevant sense, or if the meaning remained ambiguous, the reference was ignored.

Our data sources were *Masterplots* (Magill 1976) summaries read conjointly with complementary character summaries from the *Cyclopedia of Literary Characters* (Sobczak 1998). *Masterplots* is a collection of 2,010 summaries of the most prominent works in the Western literary tradition, from Homer and Hesiod through the middle 1970’s (it also includes a smattering of non-Western works that have been assimilated into the Western tradition). Summaries average 1,200-1,500 words in length. We complemented the plot summaries with character summaries from the *Cyclopedia of Literary Characters*. The *Cyclopedia* provided more detailed information on individual characters with longer summaries (approximately 200 words) for the main characters. Together, *Masterplots* and the

Cyclopedia represent something resembling a consensus on what works constitute the core of the Western literary tradition).

Using summaries allowed us to derive basic information about a large number of literary works in a reasonable amount of time. Of course, using summaries also came with significant costs. Foremost among these was the loss of valuable information that occurs when a complex literary work is reduced to a 1,200-1,500 word gloss or a complicated character is described in 200 words. It is doubtlessly true that many of the summaries we read omitted details about characters that were relevant to our study. We accept this as a shortcoming, and stress that our findings should be considered provisional until buttressed by different studies. However, given the relative simplicity of our question (we only required information on a character's agonistic status and its level of attractiveness) we thought it unlikely that this issue would significantly and systematically bias our results in one direction or another. It is also likely that information in the summaries was coloured by the personal idiosyncrasies of summary authors. However, it should be noted that the *Masterplots* and *Cyclopedia* summaries are not the products of one individual's vision: the choice of works in these references was determined collaboratively and the actual summaries were contributed by scores of English faculties from dozens of universities (see *Masterplots* 1976: v-vii). That the data sources are products of a massively collaborative effort inoculates the study, in some measure, against the possibility of inept or inadequate summarisation drastically distorting our information.

We generated a random sample of 240 *Masterplots* summaries (for a list of all sampled works see Appendix) and paired them with complementary character summaries from the *Cyclopedia of Literary Characters*. However, a problem emerged with our specific sampling method. Coders were each given different alphabetised lists of summaries and instructed to choose every twentieth work on the list. Once they reached the end of the list, they were simply to repeat the process, however, this time, they were to count from the *second* work on the list. In this way, each coder was to select his or her twenty literary works. As can be seen from the resulting list, this method of randomisation introduced a decidedly non-random element into our sample. Specifically, because the initial lists were alphabetised, this method resulted in unintended clusters of works by certain authors. For example, our list happens to exclude Emily Brönte's *Wuthering Heights*, while including several works by James Barrie and other rarely read writers. While we cannot see how this fact could bias results for or against the hypothesis, we now recognise that our selection procedures were obviously non-optimal.

Coders (12 undergraduate students from St Lawrence University, 7 male and 5 female) answered the following question about all main male and

main female protagonists and antagonists (multiple protagonists and/or antagonists could be chosen from the same work): Is the character portrayed as (1) physically attractive, (2) physically unattractive, (3) average, or (4) is there no information on this topic? Crucially, coders were kept blind to the hypothesis under test to guard against the possibility that they might consciously or unconsciously seek out or manufacture evidence that confirmed it (in content analysis jargon, the coders were naive). In total, sufficiently clear information on *both* agonistic status *and* attractiveness was available for 179 different characters: 45 male protagonists, 35 male antagonists, 73 female protagonists, and 26 female antagonists. Unfortunately, the fact that female authors were starkly under-represented in our sample (8% of total) meant that we could not conduct statistically meaningful analyses of attractiveness patterns according to authors' gender.

Finally, for the purposes of this study we assumed psychological equivalence between literary characters and real human beings. In actual fact, the relationship between the representation of behaviour and psychology in literary works and behaviour and psychology in real life is a subject that awaits systematic exploration. But, for the present, it seemed safe to provisionally adopt the traditional assumption of literary scholars that literary works reflect both general aspects of human behaviour and particular aspects of given cultures, despite the fact that they also reflect the personal idiosyncrasies of writers. More specifically, if there is a PAS in a given culture or cultures, it is reasonable to anticipate that its traces will be detectable in a reasonably large and representative sample of that culture's classic literature.

Reliability

The main challenge in content analysis utilising multiple coders is to maximise inter-coder reliability – the frequency with which different researchers code the same variable, in the same text, in exactly the same way. Content analysis of literary works presents special challenges to reliability assessment due to the length and complexity of the works (even when read in summary form) and the time-consuming nature of the coding. We undertook the following conventional measures to promote inter-coder reliability: we developed a coding question that was as uncomplicated as possible, we devised simple coding instructions, and we prepared definitions of all potentially ambiguous terms. We also assessed inter-coder reliability in a formal test. Most content analysis practitioners strive for reliability rates of 80% or better, and consider 70% to be the minimum level of adequacy (Krippendorff 1980; Neuendorf 2002; Weber 1990). All twelve

coders coded the same ten characters (five male, five female) from five randomly selected sets of plot/character summaries; inter-coder reliability in identifying protagonist-antagonist status was 78% while reliability in identifying attractiveness was 89%. These rates of reliability were nearly identical to those produced by 15 coders in an earlier study, who applied exactly the same coding question to a sample of world folktales (see Gottschall et al. 2004). This approach, in which reliability ratings are established prior to actual coding rather than having multiple coders read and code all or some fraction of the different works, is well established in content analysis and is discussed as an option in books on the subject. The advantage of this approach is that it allows the compilation of larger data sets, though some analysts argue that it does so at the cost of some degree of precision in reliability reporting (for discussion of different methods of reliability testing see Krippendorff 1980; Neuendorf 2002; Weber 1990).

We acknowledge that the relatively small number of characters examined in our reliability assessment will affect confidence in the reliability of our overall results. However, confidence in our results should also be influenced, in a positive way, by the fact that 12 naive coders, working independently, identified the same general patterns in their individual collections. While the degree of confidence generated by this fact cannot be expressed quantitatively, such regularity would not be anticipated in the output of unreliable coders.

Results

In the sample as a whole, protagonists were about 50% more likely to be portrayed as physically attractive than were antagonists. This pattern was also evident in each of three historical eras: the twentieth century, the nineteenth century, and centuries prior to the nineteenth. All protagonist-antagonist differences were statistically significant¹ except for in the twentieth-century sample where, although results fell short of the conventional significance threshold ($P=.07$), they still pointed in the predicted direction.

-
1. Significance levels are conventionally registered at the .05 and .01 levels. A significance level of $p < .05$ means that there is less than a 5% chance that a relationship between two sets of numbers arose by chance alone. (The letter “p” stands for “probability”.) A significance level of $p < .01$ means that there is less than a 1% chance that a relationship between two sets of numbers arose by chance alone. In this study, all significance results are based on Z-tests. For an interesting and painless introduction to statistics see Rowntree (1981).

Table 1: Percentage of Characters Portrayed as Physically Attractive		
	Protagonists	Antagonists
Overall	88 (n=118)	60 (n=61)**
Pre-1800	92 (n=26)	60 (n=15)*
1800-1899	93 (n=42)	55 (n=19)**
1900-forward	82 (n=50)	63 (n=27)

N=number of identifiable protagonists or antagonists with information on attractiveness

*p<.05

**p<.01

One potentially confounding factor has not yet been discussed, and relates to the differing amounts of “face time” typically allocated to protagonists and antagonists. It seems possible that the above results could be an artefact of protagonists being, on average, described at greater length and in more depth than antagonists; it could be that narrators would describe more antagonists as attractive if they only had the time to do it.

While this is a legitimate concern, our data on the proportions of physically *unattractive* antagonists and protagonists undercut this explanation. As indicated in Table 2, despite the possibility of lesser face time, antagonists were more than three times more likely to be portrayed as physically unattractive than were protagonists. This result was significant for the sample as a whole. Results across historical periods, while tending in the same direction, are not reported due to insufficient sample sizes.

Table 2: Percentage of Characters Defined as Physically Unattractive		
	Protagonists	Antagonists
Overall	12 (n=118)	39 (n=61)**

N=number of identifiable protagonists or antagonists with information on attractiveness

*p<.05

**p<.01

Finally, striking patterns emerged in comparisons of male versus female antagonists and protagonists (see Tables 3 and 4; results across historical eras not reported due to insufficient sample sizes): there were *no* significant differences in the attractiveness of female protagonists and antagonists. Both female character types were overwhelmingly more likely to be portrayed as physically attractive when information was available (see Table 3). While female antagonists *were* somewhat more likely to be defined as unattractive, this result is consistent with the possibility of random fluctuation in a small sample of female antagonists with attractiveness information. However, compared to antagonists, male protagonists were twice as likely to be portrayed as attractive, and three times less likely to be defined as unattractive (see Table 4). In short, statistically significant differences in attractiveness linked to agonistic status *applied only to males*.

Table 3: Percentage of Male and Female Characters Defined as Physically Attractive				
	Female	Female	Male	Male
	Protagonists	Antagonists	Protagonists	Antagonists
Overall	93 (n=73)	89 (n=26)	80 (n=45)	39 (n=35)**

N=number of identifiable protagonists or antagonists with information on attractiveness

*p<.05

**p<.01

Table 4: Percentage of Male and Female Characters Defined as Physically Unattractive				
	Female	Female	Male	Male
	Protagonists	Antagonists	Protagonists	Antagonists
Overall	7 (n=73)	12 (n=26)	20 (n=45)	60 (n=35)**

N=number of identifiable protagonists or antagonists with information on attractiveness

* $p < .05$

** $p < .01$

Discussion

One of the latent premises of this study is that world literature represents a varied and virtually inexhaustible repository of information about human psychology, behaviour, cognition, and culture that can be systematically mined for research in the human sciences. We are not, of course, the first to suggest this. For instance, in 1923 the psychologist William McDougall wrote, “The wise psychologist will regard literature as a vast storehouse of information about human experience, and will not neglect to draw from it what he can (McDougall 1923: 9). Yet, placing prominent exceptions like Freud and Jung aside, the mining of literary data has rarely been a regular part of psychological inquiry (see Laszlo & Cupchick 2000; Schram & Steen 2001). Of those psychologists who *have* sought to harvest literary riches, most have relied on the qualitative methods of the humanities rather than on scientific quantification (for discussion of exceptions see Laszlo & Cupchick 2000; Schram & Steen 2001). The findings of the present study raise the intriguing possibility that, by applying quantitative methods in appropriate contexts, scholars can contribute to literary and scientific knowledge at the same time.

Implications: Scientific

This study makes three specific contributions to the scientific literature on the PAS. The first is to bolster existing research by finding evidence for the phenomenon using a novel data source and methodology. Second, evidence that the stereotype is present in the literary products of different historical eras is significant, suggesting (1) that the stereotype may not be solely determined by historical contingencies, and (2) the need to adjust widespread arguments that the PAS is produced and sustained largely by images of attractive and unattractive individuals conveyed by modern advertising and entertainment media (e.g. Eagly et al. 1991; Feingold 1992; Smith, McIntosh & Bazzini 1999). Third, and finally, the finding that the stereo-type applies only to the males in this sample is inconsistent with the main trends in the empirical literature. Some research on the PAS suggests

that it does not apply identically to males and females. For instance, being overweight seems to negatively affect men's salaries but not women's (Saporta & Halpern 2002), and attractive and unattractive female political candidates face different challenges than their male counterparts (Bowman 1984). However, the main finding of large-scale analyses of scores of studies of attractiveness prejudice has been of insignificant differences in how the PAS affects the judgement and treatment of men and women (Eagly et al. 1991; Feingold 1992; Langlois et al. 2000).

One possible explanation for the male-female differences is authorial bias: 92% of the works in our sample were written by males. Some scholars have argued that "femme fatales" are more characteristic of male-authored works (e.g. Dijkstra 1988, 1998; but see Craciun, Butler & Chandler 2002). If true, the appearance of enough "bad and beautiful" femme fatales in our sample could account, at least in part, for the absence of significant attractiveness differences between female protagonists and antagonists. We would encourage an analysis of how ratings of physical attractiveness track with protagonist-antagonist status in a larger sample of female authors.

A related explanation is based on Jankowiak and Ramsey's (2000) finding that 74 of 78 diverse cultures included folktale depictions of femme fatales, which they define as women who exploit their sexual appeal to dominate, destroy, or manipulate men. By contrast, their content analysis revealed that only 25% of the cultures featured a comparable male figure, what the authors call a "male beauty fatale". This finding suggests that we may not need to invoke the hypothesis of male authorial bias to explain the absence of PAS effects for female characters; this is because similar dynamics are evident in traditional folktale traditions which often include substantial roles for female storytellers (see Biebuyck 1978; Ong 1982: 140; Coote 1978: 269; Blackburn & Flueckiger 1989: 9).

Implications: Literary

Our contribution to literary studies consists, in part, in identifying large-scale patterns in a reasonably representative sample of Western literary characters. The finding that attractive characters are more likely to be protagonists and unattractive characters are more likely to be antagonists may not come as a great shock. However, this is nothing to apologise for. A different methodological perspective can be valuable not only for helping us discover new things, but also for helping us actually verify things that most of us take for granted. This is the principle of "methodological convergence" – the idea that investigators can be most confident in findings that are reproduced using a variety of methods. A great obstacle to the progress

of knowledge in literary studies is the essential disputability of all literary information. There are few problems or questions in literary study that are so well established that they are beyond serious debate. By contrast, in fields of the human, social, and natural sciences, quantitative studies provide a firm baseline of information against which competing claims can be evaluated and systematically eliminated from contention. In short, even if this study revealed nothing that was not previously suspected, it would still be valuable for beginning to lay down an empirical baseline concerning the portrayal of different types of literary characters across different eras.

However, the finding that PAS dynamics had little, if any, application to female characters was not known before, and it is unlikely that this result – with its intriguing connections to gender debates – could have been fully appreciated without quantification. What is the meaning of this? Could it be some artefact of male authorial bias? Are we to infer the presence of large numbers of femme fatales? Or could it be that due to their disadvantages in structural power females are generally not considered formidable enough to feature as antagonists unless they are attractive? In other words, does this result mean that, for women more than for men, beauty is power? Or could the results simply be results of random chance, or of fatal errors in the design, sampling, or coding of the study?

The answer to all of these questions is a distinct “maybe”. But, then again, “maybe not”. Viable explanations for distinct attractiveness differences between male and female antagonists are easy to conceive and propose but difficult to validate. While the possible explanations we propose may sound plausible, further study will be required not only to verify the general findings of this study, but also to test these explanations and suggest which, if any, works. For example, a follow-up study could test the male authorial bias hypothesis by simply attempting to replicate the present results in a sample comprised solely of female authors.

Conclusion

Contrary to common apprehensions, the restrained use of quantitative methods will not lead to a “reduction” (in the bad sense) of the richness of literary possibility. This study generates many more interesting questions than it answers. The unanswered questions reveal a glimpse of vast fields for research programs in quantitative literary analysis that are currently lying fallow. The broad argument of this paper is that, by judiciously applying quantitative methods in appropriate contexts literary scholars can open up new fields of research, pursue old questions past the point at which qualitative methods break down, and even make contributions to science. A

more vigorous branch of quantitative literary analysis can allow us to systematically test comfortable assumptions and reveal information that might not otherwise have been discovered, or even widely suspected. While many literary problems resolutely resist quantification (what is the source of the emotional power of Book 24 of the *Iliad*?), many others may be ripe for quantitative approaches. Careful and responsible methodologies for the distillation of text into numerical data exist in diverse social science fields (e.g. Neuendorf 2002; Weber 1990). To light on one example, a large proportion of quantitative anthropological research conducted over the last several decades relies heavily on the content analyses of ethnographical narratives (often quite “literary” ones) represented in the Standard Cross-Cultural Sample, the Outline of Cultural Materials, and the Human Relations Area Files Probability Sample. There is no reason why data of similar type and value cannot be culled from literature, and put to similarly good uses.

Quantitative methods have, over the last two centuries, gradually earned a significant role in all fields of the human sciences, even in those areas (including medicine) where “the numeric method” was initially greeted with hostility and disdain (see Stigler 1986; Porter 1986; Gottschall 2005). This is not because researchers in these fields finally realised that everything could be reduced to number after all. Rather, quantitative and qualitative methods came to coexist in these fields – though sometimes uneasily – because practitioners came to appreciate that both sets of tools were utterly essential complements. We hope that this simple illustration will have some part in inspiring a similar diversification of literary methodology.

References

- Biebuyck, Daniel P.
 1978 The African Heroic Epic. In: Oinas, Felix J. (ed.) *Heroic Epic and Saga: An Introduction to the World's Great Folk Epics*. Bloomington, IN: Indiana University Press.
- Blackburn, S. & Flueckinger, J. (eds)
 1989 *Oral Epics in India*. Berkeley: University of California Press.
- Bortolussi, M. & Dixon, P.
 2003 *Psychonarratology: Foundations for the Empirical Study of Literary Response*. Cambridge: Cambridge University Press.
- Bowman, A.
 1984 Physical Attractiveness and Electability: Looks and Votes. *Women and Politics* 4: 55-66.

- Coote, Mary
1978 Serbo-Croatian Heroic Songs. In: Oinas, Felix J. (ed.) *Heroic Epic and Saga: An Introduction to the World's Great Folk Epics*. Bloomington, IN: Indiana University Press.
- Craciun, A., Butler, M. & Chandler, James (eds)
2002 *Fatal Women of Romanticism*. Cambridge, UK: Cambridge University Press.
- Dijkstra, Bram
1988 *Idols of Perversity*. Oxford: Oxford University Press.
1998 *Evil Sisters: The Threat of Female Sexuality in Twentieth-Century Culture*. New York: Owl.
- Eagly, A.H., Ashmore, R.D. & Makhijani, M.G.
1991 What is Beautiful is Good, but ...: A Meta-Analytic Review of Research on the Physical Attractiveness Stereotype. *Psychological Bulletin* 110: 109-128.
- Feingold, A.
1992 Good-looking People Are Not What We Think. *Psychological Bulletin* 111: 304-341.
- Gottschall, Jonathan & 15 Co-Authors
2004 Patterns of Characterization in Folk Tales across Geographic Regions and Levels of Cultural Complexity: Literature as a Neglected Source of Quantitative Data. *Human Nature* 15, pp. 365-382.
2005 Quantitative Literary Study: A Modest Manifesto and Testing the Hypotheses of Feminist Fairy Tale Studies. In: Gottschall, Jonathan & Wilson, David Sloan (eds) *The Literary Animal: Evolution and the Nature of Narrative*. Evanston, IL. Northwestern University Press.
- Jankowiak, William & Ramsey, Angela
2000 Femme Fatale and Status Fatale: A Cross-Cultural Perspective. *Cross-Cultural Research* 34: 57-69.
- Kenny, Anthony
1986 *A Stylometric Study of the New Testament*. Oxford: Oxford University Press.
- Krippendorff, K.
1980 *Content Analysis: An Introduction to Its Methodology*. Beverly Hills, CA: Sage.
- Langlois, J., Kalakanis, L., Rubenstein, A., Larson, A., Hallam, M., & Smoot, M.
2000 Maxims or Myths of Beauty?: A Meta-Analytic and Theoretical Review. *Psychological Bulletin* 126: 390-423.
- Laszlo, J. & Cupchick, C. (eds)
2003 Psychology of Literary Narratives: Studies of Identity and Conflict. *Empirical Studies in the Arts* 23, pp. 1-4.
- Martindale, C.
1990 *The Clockwork Muse: The Predictability of Artistic Change*. New York: Basic Books.
1996 Empirical Questions Deserve Empirical Answers. *Philosophy and Literature* 20: 347-361.

- Magill, F.N. (ed.)
 1976 *Masterplots: 2,010 Plot Stories and Essay Reviews from the World's Fine Literature*. Englewood Cliffs, NJ: Salem.
- McDougall, W.
 1923 *An Outline of Psychology*. London: Methuen.
- Moretti, F.
 2005 *Graphs, Maps, and Trees: Abstract Models for a Literary History*. New York: Verso.
- Neuendorf, K.
 2002 *The Content Analysis Guidebook*. Thousand Oaks, CA: Sage.
- Ong, Walter
 1982 *Orality and Literacy: The Technologizing of the World*. New York: Methuen.
- Porter, T.
 1986 *The Rise of Statistical Thinking*. Princeton, NJ: Princeton.
- Ramsey, J. & Langlois, J.
 2002 Effects of the "Beauty is Good" Stereotype on Children's Information Processing. *Journal of Experimental Child Psychology* 81: 320-340.
- Rowntree, D.
 1981 *Statistics without Tears: A Primer for Non-Mathematicians*. New York: MacMillian.
- Saporta, I. & Halpern, J.
 2002 Being Different Can Hurt: Effects of Deviation from Physical Norms on Lawyers' Salaries. *Industrial Relations* 41: 442-446.
- Schmidt, S.J.
 1982 *Foundations of an Empirical Study of Literature: Components of Basic Theory*. Hamburg, Germany: Helmut Buske Verlag.
- Schram, D. & Steen, G. (eds)
 2001 *The Psychology and Sociology of Literature: In Honor of Elrud Ibsch*. Amsterdam: John Bengamins.
 1992 Looking Back – Looking Ahead: Literary Studies: Trends in the Nineties. *Poetics* 21: 1-4.
- Sobczak, A.J. (ed.)
 1998 *Cyclopedia of Literary Characters*. Englewood Cliffs, NY: Salem.
- Stigler, S.
 1986 *The History of Statistics*. Cambridge, MA: Harvard.
- Thornhill, R.
 1998 Darwinian Aesthetics. In: Crawford, C. & Krebs, D. (eds) *Handbook of Evolutionary Psychology*. Mahwah, NJ: Lawrence Erlbaum.
- Weber, R. P.
 1990 *Basic Content Analysis*. Newbury Park, CA: Sage.

Appendix

240 works randomly selected for coding from *Masterplots* and *The Cyclopedia of Literary Characters*

Alain-Fournier	<i>The Wanderer</i>
Alegria, Ciro	<i>Broad and Alien Is the World</i>
Alighieri, Dante	<i>The Divine Comedy</i>
Anderson, Sherwood	<i>Poor White</i>
Aristophanes	<i>The Wasps</i>
Aristophanes	<i>The Thesmophoriazusae</i>
Azuela, Mariano	<i>The Underdogs</i>
Barclay, John	<i>Argenis</i>
Barrie, James M.	<i>Dear Brutus</i>
Barrie, James M.	<i>The Little Minister</i>
Barrie, James M.	<i>Peter Pan</i>
Barrie, James M.	<i>The Admirable Crichton</i>
Barrie, James M.	<i>Quality Street</i>
Beaumont, F. & Fletcher, J.	<i>Philaster</i>
Bjornson, Bjornstjerne	<i>The Fisher Maiden</i>
Blasco Ibanez, Vecente	<i>The Cabin</i>
Borrow, George Henry	<i>Lavengro</i>
Borrow, George Henry	<i>The Romany Rye</i>
Brecht, Bertolt	<i>The Private Life of the Master Race</i>
Broch, Hermann	<i>The Death of Virgil</i>
Brönte, Charlotte	<i>The Professor</i>
Bryne, Donn	<i>Destiny Bay</i>
Butler, Samuel	<i>The Way of All Flesh</i>
Byrne, Donn	<i>Hangman's House</i>
Cabell, James Branch	<i>The Rivet in Grandfather's Neck</i>
Cain, James M.	<i>The Postman Always Rings Twice</i>
Calderón de la Barca, Pedro	<i>The Mayor of Zalamea</i>
Calderón de la Barca, Pedro	<i>Life Is a Dream</i>
Camus, Albert	<i>The Plague</i>
Cary, Joyce	<i>Herself Surprised</i>
Cheever, Jon	<i>The Wapshot Scandal</i>
Coleridge, Samuel Taylor	<i>The Rime of the Ancient Manner</i>
Compton-Burnett, Ivy	<i>Mother and Son</i>
Conrad, Joseph	<i>Heart of Darkness</i>
Conrad, Joseph	<i>Almayer's Folly</i>

Conscience, Hendrik	<i>The Lion of Flanders</i>
Cooke, John Esten	<i>Surry of Eagles' Nest</i>
Cooper, James Feinmore	<i>The Pilot</i>
Cooper, James Feinmore	<i>The Redskins</i>
Corneille, Pierre	<i>Polyeucte</i>
Davis, H. L.	<i>Honey in the Horn</i>
de Goncourt, E. & de Goncourt, J.	<i>Charles Demailly</i>
Defoe, Daniel	<i>Moll Flanders</i>
Dickens, Charles	<i>Tale of Two Cities</i>
Dos Passos, John	<i>Three Soldiers</i>
Dryden, John	<i>Marriage a la Mode</i>
Dryden, John	<i>Absalom and Achitophel</i>
Duggan, Alfred	<i>Leopards and Lilies</i>
Duggan, Alfred	<i>The Lady for Ransom</i>
Dumas, Alexandre	<i>The Vicomte de Bragelonne</i>
Duun, Olav	<i>The People of Jurik</i>
Earle	<i>Eugene Aram</i>
Edgeworth, Maria	<i>Castle Rackrent</i>
Etherege, Sir George	<i>The Man of Mode</i>
Euripedes	<i>The Cyclops</i>
Euripedes	<i>Medea</i>
Farquhar, George	<i>The Beaux Stratagem</i>
Farrell, James T.	<i>Studs Lonigan</i>
Faulkner, William	<i>The Unvanquished</i>
Faulkner, William	<i>The Hamlet</i>
Faulkner, William	<i>Requiem for a Nun</i>
Faulkner, William	<i>Sanctuary</i>
Faulkner, William	<i>A Fable</i>
Faulkner, William	<i>The Sound and the Fury</i>
Fernandez de Lizardi, Jose Joaquin	<i>The Itching Parrot</i>
Fletcher, John	<i>Rule a Wife and Have a Wife</i>
Fletcher, John	<i>The Faithful Shepherdess</i>
Fogazzaro, Antonio	<i>The Patriot</i>
Forster, E. M.	<i>A Room with a View</i>
Forster, E. M.	<i>A Passage to India</i>
Fromentin, Eugene	<i>Dominique</i>
Gaboriau, Emile	<i>File No. 113</i>
Galt, John	<i>Annals of the Parish</i>
Gibran, Kahlil	<i>The Prophet</i>

ARE THE BEAUTIFUL GOOD IN WESTERN LITERATURE?: ...

Gissing, George	<i>The Private Papers of Henry Ryecroft</i>
Goethe, Johann Wolfgang von	<i>Faust</i>
Gordon, George (Lord Byron)	<i>Manfred</i>
Gourmont, Remy de	<i>A Night in the Luxembourg</i>
Greene, Graham	<i>The Power and the Glory</i>
Haggard, H. Rider	<i>King Solomon's Mine</i>
Halevy, Ludovic	<i>The Abbe Constantin</i>
Hardy, Thomas	<i>Under the Greenwood Tree</i>
Hardy, Thomas	<i>The Return of the Native</i>
Hawthorne, Nathaniel	<i>The Scarlet Letter</i>
Hawthorne, Nathaniel	<i>The Marble Faun</i>
Hemingway, Ernest	<i>A Farewell to Arms</i>
Hemingway, Ernest	<i>The Sun also Rises</i>
Hergesheimer, Joseph	<i>The Three Black Pennies</i>
Herlihy, James Leo	<i>All Fall Down</i>
Hesse, Hermann	<i>Steppenwolf</i>
Heyward, DuBose	<i>Porgy</i>
Homer	<i>The Odyssey</i>
Hope, Anthony	<i>The Prisoner of Zenda</i>
Howe, Edgar Watson	<i>Story of a Country Town</i>
Howells, William Dean	<i>The Rise of Silas Lapham</i>
Hutchinson, A. S. M.	<i>If Winter Comes</i>
Ibsen Henrik	<i>Hedda Gabler</i>
Ibsen, Henrik	<i>When We Dead Awaken</i>
Ibsen, Henrik	<i>Brand</i>
James, Henry	<i>Daisy Miller</i>
James, Henry	<i>Roderick Hudson</i>
James, Henry	<i>The Sacred Fount</i>
James, Henry	<i>The Spoils of Poynton</i>
Jeffers, Robinson	<i>Tamar</i>
Johnson, Ben	<i>Volpone</i>
Johnson, Ben	<i>Catiline</i>
Johnson, Ben	<i>Bartholomew Fair</i>
Johnson, Ben	<i>Every Man out of His Humour</i>
Johnson, Ben	<i>Every Man in His Humour</i>
Joyce, James	<i>Exiles</i>
Kazantzakis, Nikos	<i>Freedom or Death</i>
Kingsley, Amis	<i>That Uncertain Feeling</i>

Kleist, Heinrich von	<i>Prince of Homburg</i>
Kyd, Thomas	<i>The Spanish Tragedy</i>
La Fayette, Madame Marie de	<i>The Princess of Cleves</i>
Lawrence, D. H.	<i>Sons and Lovers</i>
Le Fanu, Joseph Sheridan	<i>Uncle Silas</i>
Le Fanu, Joseph Sheridan	<i>The House by the Churchyard</i>
Le Sage, Alain René	<i>Le Diable Boiteux</i>
Lessing, Gotthold Ephraim	<i>Nathan the Wise</i>
Lessing, Gotthold Ephraim	<i>Minna Von Barnhelm</i>
Lewis, Sinclair	<i>Cass Timberlane</i>
Lie, Jonas	<i>The Family at Gilje</i>
London, Jam	<i>The Sea Wolf</i>
Longus	<i>Daphnis and Chloe</i>
Lonnrot, Elias	<i>The Kalevala</i>
Loti, Pierre	<i>An Iceland Fisherman</i>
Lover, Samuel	<i>Handy Andy</i>
Lyly, John	<i>Campaspe</i>
Lyly, John	<i>Euphues, The Anatomy of Wit</i>
Maeterlinck, Maurice	<i>Pelleas and Melisande</i>
Mann, Thomas	<i>Doctor Faustus</i>
Mann, Thomas	<i>Buddenbrooks</i>
Massinger, Philip	<i>A New Way to Pay Old Debts</i>
Maugham, W. Somerset	<i>Liza of Lambeth</i>
Maugham, W. Somerset	<i>Cakes and Ale</i>
Melville, Herman	<i>Typee</i>
Meredith, George	<i>Evan Harrington</i>
Merejkowski, Dmitri	<i>The Romance of Leonardo Da Vinci</i>
Miller, Arthur	<i>Death of a Salesman</i>
Milton, John	<i>Paradise Lost</i>
Molière, Jean Baptiste Poquelin	<i>The Miser</i>
Moore, Thomas	<i>Lalla Rookh</i>
Murdoch, Iris	<i>The Bell</i>
Nabokov, Vladimir	<i>Pale Fire</i>
O'Brien, Kate	<i>The Last of Summer</i>
O'Hara, John	<i>Appointment in Samara</i>
O'Neill, Eugene	<i>Strange Interlude</i>
Otway, Thomas	<i>The Orphan</i>
Page, Thomas Nelson	<i>Marse Chan</i>

ARE THE BEAUTIFUL GOOD IN WESTERN LITERATURE?: ...

Paton, Alan	<i>Cry, the Beloved Country</i>
Peele, George	<i>The Old Wives' Tale</i>
Perez Galdos, Benito	<i>Saragossa</i>
Perez Galdos, Benito	<i>Angel Guerra</i>
Perez Galdos, Benito	<i>Fortunata and Jacenta</i>
Petronius	<i>The Satyricon</i>
Pilnyak, Boris	<i>The Naked Year</i>
Pinero, Arthur Wing	<i>Mid-Channel</i>
Plautus	<i>The Captives</i>
Plautus	<i>Amphitryon</i>
Poe, Edgar A.	<i>Ligeia</i>
Pontoppidan, Henrik	<i>The Promised Land</i>
Porter, Jane	<i>The Scottish Chiefs</i>
Porter, Katherine Anne	<i>Old Mortality</i>
Powell, Anthony	<i>A Dance to the Music of Time</i>
Pushkin, Alexander	<i>Eugene Onegin</i>
Racine, Jean Baptiste	<i>Mithridate</i>
Remarque, Erich Maria	<i>All Quiet on the Western Front</i>
Rice, Elmer	<i>Street Scene</i>
Rulfo, Juan	<i>Pedro Paramo</i>
Saint-Beuve, Charles Augustin	<i>Volupté</i>
Sandburg, Carl	<i>Remembrance Rock</i>
Santayana, Geoge	<i>The Last Puritan</i>
Scott, Sir Walter	<i>The Lady of the Lake</i>
Scott, Sir Walter	<i>The Heart of Midlothian</i>
Scott, Sir Walter	<i>St. Ronan's Well</i>
Scott, Sir Walter	<i>The Fair Maid of Perth</i>
Shakespeare, William	<i>Two Gentlemen of Verona</i>
Shakespeare, William	<i>King John</i>
Shakespeare, William	<i>Richard the Third</i>
Shakespeare, William	<i>Much Ado about Nothing</i>
Shakespeare, William	<i>The Tempest</i>
Shelley, Percy Bysshe	<i>The Cenci</i>
Shelley, Percy Bysshe	<i>Prometheus Unbound</i>
Sherriff, Robert C.	<i>Journey's End</i>
Sholokhov, Mikhail	<i>And Quiet Flows the Don</i>
Singer, Israel Joshua	<i>The Brothers Ashkenazi</i>
Spark, Muriel	<i>The Bachelors</i>
Stendhal	<i>The Red and the Black</i>

Stephens, James	<i>Deirdre</i>
Stevenson, Robert Louis	<i>Kidnapped</i>
Stingmaster, Elsie	<i>A High Wind Rising</i>
Suckow, Ruth	<i>The Folks</i>
Synge, John Millington	<i>The Playboy of the Western World</i>
Syron, William	<i>Lie down in Darkness</i>
Tasso, Torquato	<i>Jerusalem Delivered</i>
Terence	<i>Phormio</i>
Tourgee, Albion	<i>A Fool's Errand</i>
Trollope, Anthony	<i>Doctor Thorne</i>
Trollope, Anthony	<i>Barchester Towers</i>
Trollope, Anthony	<i>Framley Parsonage</i>
Turgenev, Ivan	<i>A Month in the Country</i>
Twain, Mark	<i>The Prince and the Pauper</i>
Undset, Sigrid	<i>The Axe</i>
Unknown	<i>Havelok the Dane</i>
Unknown	<i>Hercules and His 12 Labors</i>
Unknown	<i>Jason and the Golden Fleece</i>
Unknown	<i>Reynard the Fox</i>
Unknown	<i>Aucassin and Nicolette</i>
Upton, Sinclair	<i>The Jungle</i>
Wallant, Edward Lewis	<i>The Tenants of Moonbloom</i>
Waugh, Evelyn	<i>Decline and Fall</i>
Webb, Mary	<i>Precious Bane</i>
Wells, H. G.	<i>Kipps</i>
Wells, H. G.	<i>The Invisible Man</i>
Welty, Eudora	<i>Delta Wedding</i>
Westcott, Edward Noyes	<i>David Harum</i>
Wescott, Glenway	<i>The Pilgrim Hawk</i>
Whyte-Melville, George J.	<i>Digby Grand</i>
Wilde, Oscar	<i>The Importance of Being Earnest</i>
Wilder, Thornton	<i>Heaven's My Destination</i>
Wilder, Thornton	<i>The Bridge of San Luis Rey</i>
Wilder, Thornton	<i>Our Town</i>
Williams, William Carlos	<i>Paterson</i>
Wilster, Owen	<i>The Virginian</i>
Woolf, Virginia	<i>Orlando</i>
Wylie, Elinor	<i>The Venetian Glass Nephew</i>
Zorrilla, José	<i>Don Juan Tenorio</i>

ARE THE BEAUTIFUL GOOD IN WESTERN LITERATURE?: ...