“Toupee or not Toupee?” Sic (et Non). Social perceptions of Male Pattern Baldness (Androgenetic Alopecia)

Ronald Henss

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.

Abstract

13 men aged between their early 30s and early 70s with full-blown male pattern baldness (male androgenetic alopecia) were photographed once bald and once with a high-quality, custom-made toupee. In an online experiment, one sample rated the photos according to attractiveness, another according to self-assurance and another according to health. The attractiveness ratings were extraordinarily low. The men appeared considerably more attractive with a toupee than with a bald head (effect size Cohen’s d = 0.67). The women’s attractiveness ratings were more negative than those of the male judges (d = 0.58). When the men wore their toupee, they were rated higher by older judges than by younger ones. The men appeared somewhat healthier with a toupee than with a bald head (d = 0.14). Overall, neither the hair status nor the sex of the judges played a role in the assessment of self-assurance. A special feature of the study is the analysis at the level of the individual stimulus persons. Under both hair conditions, each individual received higher attractiveness ratings from men than from women. In terms of self-assurance, some were rated more favorably by women and others by men. In terms of health, men and women only differed for a few targets. In line with the attractiveness stereotype „What is beautiful is good“, more attractive men were judged to be more self-assured and healthier.

Keywords: Male Pattern Baldness, Androgenetic Alopecia, Social Perception, Face Perception, Toupee, Cranial Hair,

Physical Attractiveness, Self-assurance, Health, Attractiveness Stereotype, Online Experiment.

Highlights

- When wearing their toupee, men are considerably more attractive than when bald
- Women give much lower attractiveness ratings than men
- Attractiveness of stimulus persons is strongly devalued
- Men appear slightly healthier with a full head of hair than with a bald head
- Hair status has no overall influence on assessment of self-assurance
- Analysis at the level of individual stimulus persons provides a differentiated picture

Introduction

The ingenious question that opens the title of our paper is borrowed from Wade, Fisher and Burch (2021), whose borrowing from Shakespeare is evident. Our answer is borrowed from Peter Abaelard, arguably one of the most ingenious and influential scholars of the Middle Ages. In his work *Sic et Non* (*Yes and No*), he juxtaposed contradictory theological statements and, with this work, raised dialectics and logic to a higher level in the discourse of scholasticism. In this article, we do not deal with theological and philosophical questions any more than we do with Shakespeare. Instead, we deal with a tangible worldly problem of men, which may seem far-fetched to some, but which in no way is.

The focus of our interest is on *hereditary hair loss in men*, also known as *male pattern baldness* or, in medical terminology, *male androgenetic alopecia*. This specific form of hair loss affects a majority of men of European descent to a greater or lesser extent in the course of their lives. Our study is not about the perspective of those affected, but about the impression of the amount of hair on strangers, particularly with regard to physical attractiveness. It is about social perception, about the proverbial eye of the beholder.

Almost a quarter of a century ago, we outlined the theoretical background and the state of empirical research in a review (Henss, 2001). Readers who are less familiar with the subject area are encouraged to have a look at this paper, which is freely available on the Internet.[1]

We keep hearing the warning that one should not judge people by their outward appearance – Don't judge a book by its cover – but this advice is as quixotic as it is infeasible. In many situations in everyday life, we cannot afford to wait until we have discovered the inner values of our fellow human beings, and sometimes there is no other information available to us than their physical appearance. Consequently, in the course of evolution, psychological mechanisms have developed that lead to automatic, instantaneous and unconscious assessments based on external characteristics. Among the features that serve as key cues, the face and hair play a paramount role.

Our perceptual system is so highly specialized that we can differentiate between thousands of faces, some of which differ only in minute details, and different configurations evoke different impressions. The hair, which in a sense constitutes the frame of the face, can modify these impressions to a substantial degree.

Physiognomic features change only slowly and unnoticeably over the course of a lifetime; and the face can essentially only be changed through facial expressions, cosmetic means or medical techniques. Hair is quite different. As the frame of the face and the uppermost point of the human body, it is inevitably the focus of attention, it already exhibits a great deal of inherent variability and, as it is dead matter, it can be effortlessly altered in countless ways like no other part of the body. This results in an endless wealth of variations that enable powerful social signaling functions. In our review, we have described these unique properties of human hair in detail under the headings of Visual Prominence, Variability, Manipulability and Display Functions.

From the broad spectrum of highly interesting questions relating to hair, we restrict ourselves to an area which, in contrast to the intensive medical research, receives remarkably little attention in psychological research, although it affects not only an exotic minority, but, as mentioned, to a greater or lesser extent the majority of men of European descent.

Male androgenetic alopecia undoubtedly has a strong polygenetic foundation. Genome-wide association studies (GWAS) have identified more than three hundred loci that may be associated with this form of hair loss (Henne, Nöthen and Heilmann-Heimbach, 2023, p. 5). The data on prevalence are somewhat inconsistent, but there is no doubt that Whites are much more severely affected than Blacks or East Asians, for example. Henne et al. report „a lifetime prevalence of ~80% in European men“ (p. 3). However, this figure refers to a broad spectrum of different manifestations. In our study, we restrict ourselves to the final stage, which of course is not reached by all those affected, but in some cases is already pronounced in 30-year-olds. It should be noted that androgenetic alopecia must not be equated with a fully shaven head. Even in the final stage, a horseshoe-shaped fringe of hair remains and therefore androgenetic baldness and a completely shaved head can give very different impressions.[2]

Of course, baldness caused by androgenesis can also evoke quite different impressions than a full head of hair. The contrast between full hair and full-blown androgenetic alopecia is illustrated in Figures 1 and 2. The images were generated by artificial intelligence. They were not used in our study. Our stimulus persons are real men of flesh and blood.

Figure 1. Full hair versus androgenetic alopecia. Created by Artificial Intelligence.

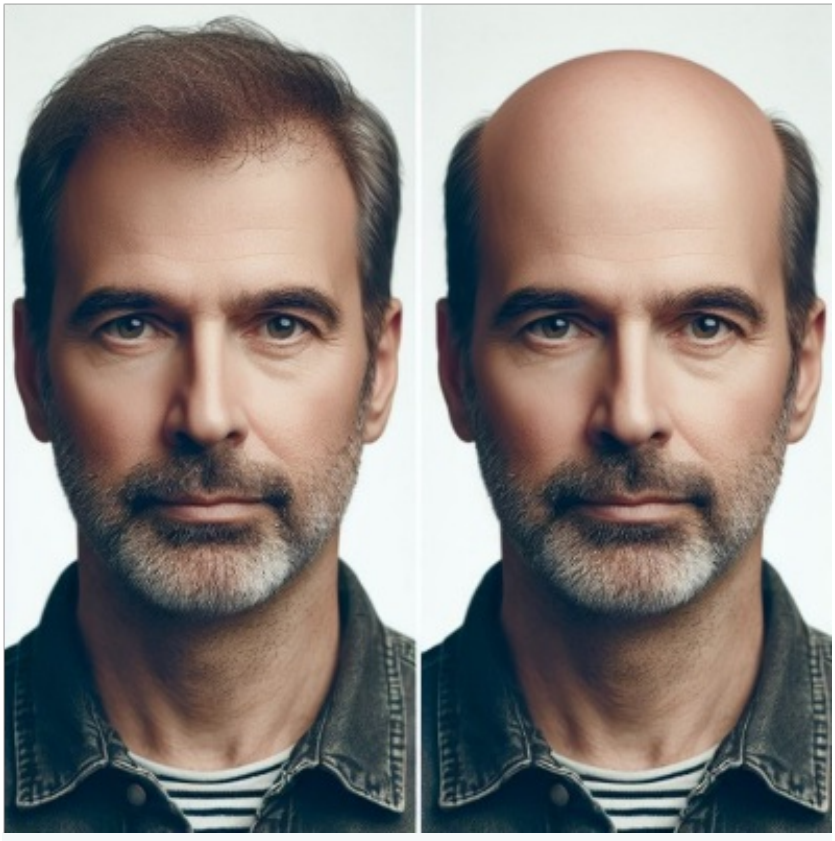

Figure 2. Full hair versus androgenetic alopecia. Created by Artificial Intelligence.

The contrast was noticed by humans thousands of years ago. And because they know that a bald head has negative effects in many ways, men have been fighting against the loss of hair for thousands of years. The oldest written evidence we know of can be found in the Papyrus Ebers, a well-preserved twenty meter long Egyptian scroll from the 16th century BC. It contains formulations for various ailments, including „a number of recommended cures for Egyptians suffering from hair loss. Suggestions include a mixture of fats from a hippopotamus, crocodile, tomcat, snake and ibex; porcupine hair boiled in water and applied to the scalp for four days; and the leg of a female greyhound sautéed in oil with the hoof of a donkey. Should none of these lead to the hoped for renewal of hair growth, both male and female royals in ancient Egypt were known to wear wigs and fake beards" (Homan, 2019, S. 24).[3]

Since then, mankind has made some progress in the fight against hair loss.[4] Thanks to the fantastic advances in medicine, androgenetic alopecia can now be treated in a causal way. The dream of constantly regrowing hair has not yet been fulfilled, but we now have effective pharmaceutical agents that can slow down or even stop further loss. In addition, there are modern medical techniques to reforest bald areas. And the millennia-old arsenal of camouflage methods has been enriched by top-quality toupees that are almost indistinguishable from natural hair.

It is sometimes criticized that male hair loss is a multi-billion dollar business and that the subject is heavily medicalized.[5] Both arguments are valid. However, it should be borne in mind that the fight against hair loss has always been medicalized, albeit at a very primitive and ineffective level, and that those who could afford it spared no expense and that there have always been players for whom the fight against hair loss has brought a sizable income.

Androgenetic alopecia has biological causes and causal treatment requires medical methods. However, it is not primarily a medical problem, but a psychological one. Hair loss does not hurt, it is not a disability, it is not contagious and it does not cause material damage. It is a problem for those affected because they know that, despite all pseudo-moral teachings, the outward appearance is a very important factor and that baldness has adverse effects on others.

It is not only that the amount of hair correlates negatively with age and that bald men appear older than they actually are. The negative effects primarily concern physical attractiveness, which is by no means only a matter for women, but also for men. In addition, bald men are also judged less favorably than men with a full head of hair with regard to numerous other personality traits.[6] Furthermore, hair status not only affects social perception, hair loss can also have tangible negative consequences in real life.

If fellow human beings were indifferent to outward appearance, there would be no cause for concern and hair status would not be an issue. There are probably only a few people affected who are completely unconcerned about their hair status. Most of those affected can cope well with progressive hair loss, but there are also many for whom it is a problem and for some it is even a serious problem.[7]

As the root of all these problems lies in the psychological impact of hair loss on others, one would expect this topic to receive a great deal of attention in psychological research. But this is not the case. In our review, we outlined the state of research up to the year 2000. We pointed out that only very few empirical studies were available and that some of these had considerable methodological shortcomings, primarily relating to the stimulus material. With regard to this aspect, we

would like to briefly address three recent studies.

The first is the study by Wade, Fisher and Burch (2021), from which we borrowed our title question „Toupee or not Toupee?“. The abstract starts with the statement „The question of whether or not cranial hair affects perceptions of attractiveness, personality, career success, and other traits related to fitness for men in two populations was investigated in two experiments“. However, a glance at the experimental design makes it clear that the aim of the study cannot be achieved in any way. The stimulus subjects were *a single* white man and *a single* black man. Both were bald and both were photographed once with a bald head and once with a toupee. So there were only four stimuli, one bald and one toupee photo of *a single* white man and one bald and one toupee photo of*a single* black man. With such an experimental set-up, it is a priori clear that neither any generalization to the population of Whites nor to the population of Blacks is possible, let alone a meaningful comparison of the two races. All of this is entirely independent of what the participants' assessments may reveal. And no non-trivial generalization is possible about the effect of baldness and full hair.

In stark contrast to this is the study by Kranz, Nadarevic and Erdfelder (2019). In a theoretically and methodologically sophisticated study, they investigated the implicit and explicit activation of the attractiveness stereotype in the sense of „What is beautiful is good“ (Dion, Berscheid and Walster, 1972; Eagly et al., 1991; Henss, 1992, 1998). Images of 10 „men“ were used as stimulus material, one with a bald head and one with a full head of hair. The word men is placed in quotation marks because the authors unfortunately used software-generated and retouched faces as models.[8] Thus, the study allows statements about faces that look more like ghosts than flesh-and-blood humans.

Exemplary pictorial material was used by Chan et al. (2019). They investigated the impact of masculine facial features of Swiss politicians that were unknown to the predominantly Australian judges, as well as the relationship between masculine features and the actual behavior of the politicians. One of the cues was the amount of hair, with baldness considered an indicator of masculinity. High-quality standardized photographs of „52 male Swiss Members of Parliament (Upper House) who were politically active between 2007 and 2014“ (p. 3) served as stimulus material.[9] Here, the abundance of hair was not varied experimentally in a systematic way, but in a quasi-experiment with pre-existing stimulus material. This is not a serious problem. If the number of stimulus persons is sufficiently large, this approach also enables far-reaching conclusions to be drawn, as the authors make well clear. However, even more substantial conclusions can be drawn if the one and the same persons are presented once with a bald head and once with a full head of hair. This is what we did in our study.

Our stimulus persons show the typical appearance of full-blown androgenetic alopecia; and they own a custom-made, high-quality toupee. Under standardized conditions, a professional photographer took one photograph with a bald head and one with a toupee. The pictures were evaluated by large samples in an online experiment. One sample assessed *attractiveness*, another *self-assurance* and another one *health*. Our focus is on physical attractiveness. Health was considered because it is an important component of mate value (on mate value theory, see for example Buss, 2016[4]; Henss, 1992, 1998), and because androgenetic alopecia is often seen as a disease although it may better be regarded as a psychological problem. Self-assurance was included because it is associated with masculinity, which is also a component of mate value. In addition, some of our studies indicate that when assessing the faces of strangers, traits such

as self-assured, self-confident, independent, self-reliant, secure, superior, independent, self-satisfied can constitute an independent factor that stands out from the Big Five or other conceptions of personality research (Henss, 1997, 1998). A special feature of our study is that we do not only compare baldness versus full hair at the group level, as is usually the case. We also look at the level of the individuals. Although it seems self-evident that the difference between baldness and a full head of hair does not generate the same effect on the observer for all target persons, this aspect is not usually addressed at all.

Method

Stimulus persons and material

The stimulus subjects were 13 men in their early 30s to early 70s. They are natural bald people with a characteristically pronounced pattern of androgenetic alopecia. In their everyday life, they wear a toupee that is not an off-the-peg product. The hairpieces are custom-made by a renowned toupee specialist and are of such high quality that they blend in seamlessly and almost imperceptibly with the natural hairline in terms of hair texture, color, density and mottling. As they are individually crafted and professionally tailored to the customer's preferences, they perfectly simulate the look of the natural hairstyle. In everyday life, it would be quite difficult to distinguish them from natural hair. In photographs, this is impossible. The stimulus material was photographs taken under standardized conditions by a professional photographer. The photos were full-face portraits with a neutral facial expression and looking straight into the camera. Each person was first photographed without and then with his toupee.

Procedure

The study was conducted as an online experiment in a German-language version. Participants were recruited via links on our homepage, which was well frequented due to numerous previous experiments. Participation was voluntary, non-binding and anonymous.

The participants were shown 13 photos one after the other, which they were asked to rate according to a fixed criterion on a 9-point bipolar rating scale. The study was conducted in three consecutive time intervals. In experiment 1, the stimulus persons were to be judged according to their *Attractiveness* (unattractive – attractive), in experiment 2 according to their *Self-assurance* (insecure – self-assured) and in experiment 3 according to their*Health* (ill – healthy).[10] The order of the stimulus persons was determined at random, as was the hair status (Bald or Toupee) in each individual case. Thus, the participants saw different combinations in a completely randomized experiment.

Following the face assessments, the participants were asked to assess their own current mood state using the PANAS (Positive and Negative Affect Scale) by Watson, Clark and Tellegen (1988). The questionnaire comprises 20 adjectives on various aspects of mood, which are graded on a 5-point rating scale. Ten items each constitute a factor, which is referred to as Positive Affect or Negative Affect. Contrary to what the labels suggest, these are not opposing poles of a common

dimension, but *orthogonal* factors. This has been confirmed time and again in countless studies in various countries. The self-assessment of mood serves a different purpose. We will only deal with this in passing; a detailed analysis will be reported elsewhere.

Hypotheses

With regard to the *attractiveness* ratings, we have two clear hypotheses. Firstly, we expect a pronounced advantage for the toupee; and secondly, we expect women to give considerably less favorable ratings than men. Based on experience from numerous previous studies on face perception, we also expect the overall level of evaluation to be very low.

For the criteria of self-assurance and health, considerably weaker effects are expected.

In the case of *health*, it can be assumed that men with a toupee are rated more favorably when they wear their toupee.

No hypotheses are specified for *self-assurance*, as the picture is contradictory. On the one hand, greater insecurity could be expected, as men are well aware that they are less attractive in the eyes of others; on the other hand, however, baldness may be seen as a signal of masculinity and dominance, which can go hand in hand with greater self-assurance.

Participants

Table 1 shows the number of participants and statistical parameters of their age (mean, standard deviation, skewness, kurtosis).

Table 1. *Number and age of participants. Mean, standard deviation, skewness, kurtosis.*

		N	M	sd	Skewness	Kurtosis
Attractiveness	**Males**	89	31,8	11,8	0,85	0,29
	Females	276	25,9	8,6	1,53	2,31
Self-assurance	**Males**	657	30,7	11,5	0,95	0,38
	Females	1295	27,0	10,0	1,31	1,31
Health	**Males**	389	30,1	11,5	0,87	-0,10
	Females	1019	25,8	9,3	1,37	1,44

In all cases, the women are markedly younger than the men. For the attractiveness ratings, the mean difference is 5.9 years, for self-assurance 3.7 and for health 4.3. For the median, the difference is 6, 3 and 4 years respectively. The range is similar in each case and extends from 16 to the end of 60, in one case up to 77 years. Such a pattern is common for self-selection in our online experiments. It is also common for the majority of participants to come from academic and school backgrounds.

The first data column shows that our samples are unusually large. This means that our tests have a very high test power.

This applies in particular to the assessments of self-assurance and health. Here, even very small differences can become statistically significant. We will therefore not only report the statistical significance, but also the effect size, as measured by Cohen's d. In a comparison of two samples, d corresponds to the difference between the mean values divided by the standard deviation (pooled in the case of unequal samples). There are different opinions on the interpretation of this measure. One common rule of thumb is: d = 0.1 weak, d = 0.3 medium, d = 0.5 strong; another is: d = 0.2 weak, d = 0.5 medium, d = 0.8 strong. We will follow the first variant, as this is more in line with the actual situation in social science studies. In the discussion, we will propose a different view on Cohen's d.

Results

In the following, we first look at the global level. Here, the men are grouped together when they are bald as well as when they wear their toupee. Then we look at the individual stimulus persons. Finally, we briefly look at the assessment of the current mood of the participants.

Global perspective

The dependent variables at the global level are the average ratings of the men, once bald and once with a toupee. It should be remembered here that the raters saw each man in only one variant and had to judge different combinations and sequences of stimulus person and hair status.

For each of the three criteria, the data were analyzed in a 2 x 2 analysis of variance with the hair status as a within-subjects factor and the sex of the judges as a between-subjects factor. In each case, both the variance homogeneity and a normal distribution of the standardized residuals were fulfilled (Levene's test and Q-Q plot).

Figure 3 shows the means for the three criteria broken down by hair status and sex. On the Y-axis, we deliberately plotted the full scale width in order to visualize the actual relationships. It should be borne in mind that the scale does not start at 0, but at 1; and this can easily be misleading when interpreting the numerical values – a recurring problem that is all too rarely addressed.

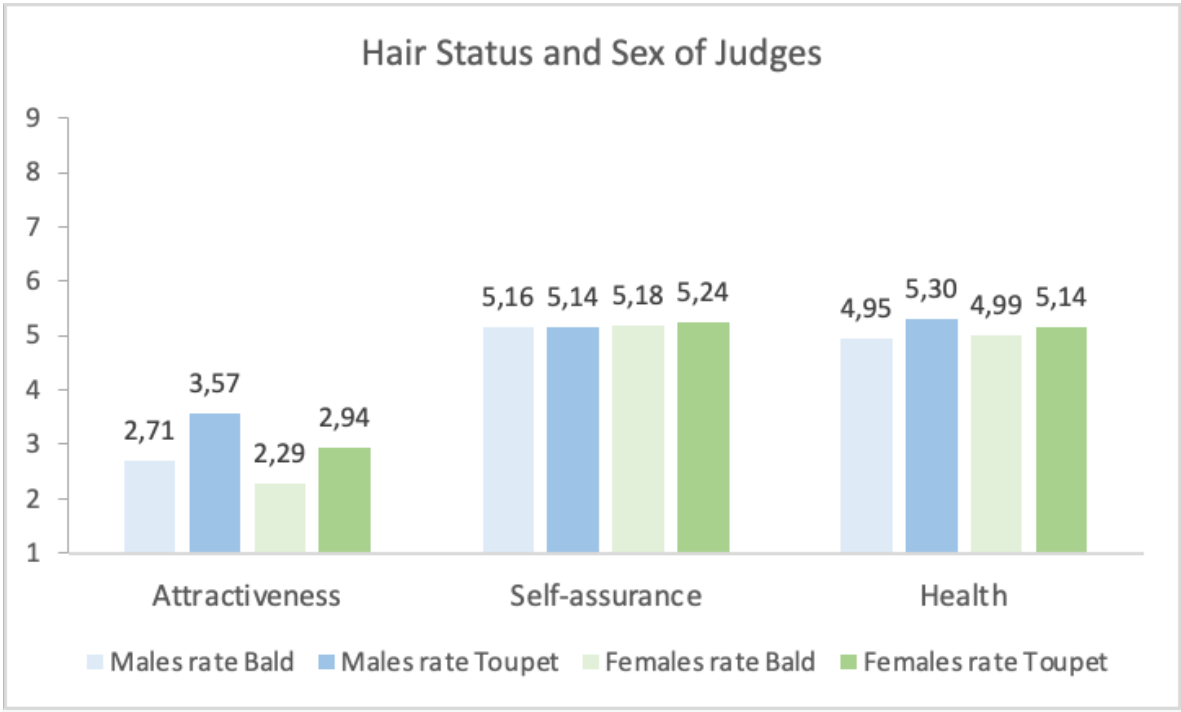

Figure 3. Mean ratings by Hair Status and Sex of Judges.

The special role of the attractiveness ratings immediately catches the eye. While the mean ratings on self-assurance and health are quite close to the center of the 9-point scale for all combinations, the attractiveness ratings are extraordinarily low. It should be emphasized that the stimulus persons were by no means a negative selection. On the contrary, these are men as one would meet them in everyday life in public spaces. Nevertheless, the attractiveness ratings are decidedly negative. We will come back to this point in detail in the discussion.

In addition, another point is immediately apparent: in the assessment of self-assurance, hair status and sex play no role and in the case of health, the differences are not great, while there are considerable differences – at a very low level – when it comes to attractiveness.

For *attractiveness*, both main effects are significant at the 0.001 per cent level and there is no interaction between the factors. When wearing their toupee, the men are rated 0.71 points higher than when bald (3.10 vs. 2.39), which corresponds to an effect size of d = 0.67. The women's judgment is even more derogatory than the men's (2.65 vs. 3.14). The effect size is 0.58.

Neither hair status nor sex makes a difference to *self-assurance*.

Hair status has a significant effect on *health* ($p < 0.001$). When the men wear a toupee, they appear 0.20 points healthier (d = 0.14). The sex of the judges has no effect on its own, but the interaction between the two factors is significant ($p = 0.021$). As can be seen in Figure 3, the men differentiate more strongly between baldness and toupee than the women. In the eyes of the men, the toupee gives a gain of 0.35 points, in the eyes of the women 0.15.

So far we have ignored one important factor, namely the *age of the judges*. However, it can be assumed that age plays a

role, and since, as seen in Table 1, age and sex of the judges are confounded, it is necessary to take their age into account as well.

In an analysis of covariance with age as a covariate, the interaction with sex is not significant, instead there is a significant interaction with hair status ($p = 0.002$). A linear regression on the age of the judges is not significant under the Bald condition ($p = .326$). When the men wear their hairpiece, however, the age of the judges has a significant impact ($p < .001$; $R = .298$; Adjusted $R^2 = .084$; Intercept = 2.855; Slope = 0.0225).

To illustrate the effect of age, we formed the four age groups [16 - 20], [21 - 26], [27 - 36] and [37+]. The intervals were chosen so that a sufficient number of men were represented in each group. In an analysis of variance, hair status and sex are significant ($p < .001$), but age group is not ($p = .254$). In addition, there is a significant interaction hair status x age group ($p = .011$). All other interactions are insignificant. The mean scores are depicted in Figure 4. Here, too, the full scale width is plotted on the Y-axis in order to visualize the degrading evaluation.

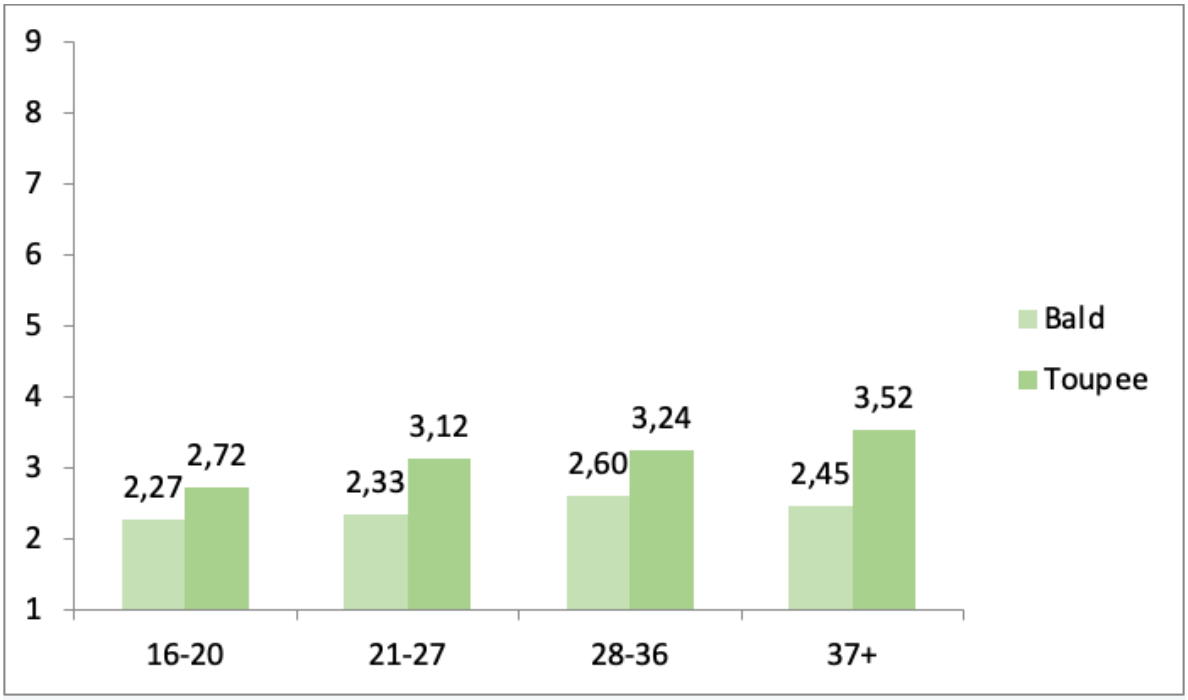

Figure 4. Mean scores. Hair Status by Age of Judges.

When the men are bald, they are rated more favorably by the older judges than by the younger ones, but the trend is not monotonic and the maximum difference is 0.33 points. When the men are seen with a full head of hair, the older the judges are, the better the men are rated, and the difference between the youngest and oldest respondents is 0.80 points. This is a huge difference.

A similar picture emerges in the assessment of*health*, but the impact of age is considerably weaker. In the assessment of *self-assurance*, the age of the judges plays no role at all.

Focus on individual stimulus persons

The analysis at the group level reveals the general trend; however, there are some large differences between the individuals. Table 2 shows the mean scores for the individuals, ranked in descending order according to their attractiveness under their natural condition, i.e. when bald.

Table 2. *Mean scores of individual stimulus persons.*

	Attractiveness		Self-assurance		Health	
No.	**Bald**	**Toupee**	**Bald**	**Toupee**	**Bald**	**Toupee**
1	3.35	4.43	6.72	7.19	6.63	7.49
2	2.95	4.40	5.99	6.42	5.59	6.23
3	2.87	3.08	5.47	4.35	4.85	4.32
4	2.83	3.36	5.39	4.49	5.76	6.01
5	2.62	2.17	5.26	5.46	4.02	3.85
6	2.57	3.57	4.82	5.08	6.02	6.65
7	2.37	2.66	4.67	4.60	5.20	5.00
8	2.32	3.04	5.07	5.22	4.81	5.05
9	2.27	3.92	5.14	5.58	4.72	5.79
10	2.16	2.51	5.31	5.28	5.24	5.35
11	1.84	2.58	4.94	5.56	4.18	4.69
12	1.73	2.10	4.02	4.50	3.57	3.62
13	1.72	2.17	4.40	3.82	3.89	3.59

Due to the wealth of data, the table may appear confusing at first glance, but the situation is actually straightforward.

The two top performers remain at the top and the two bottom performers remain at the bottom and, with a few exceptions, the variations in between are quite small. Table 3 documents the stability of the basic pattern by means of the correlation between the columns.

Table 3. *Correlation over target persons.*

	A Bald		A Toupee		S Bald		S Toupee		H Bald	
A Toupee	.74	**								
S Bald	.87	***	.76	**						
S Toupee	.56	*	.69	**	.79	**				
H Bald	.79	**	.79	**	.72	**	.53			
H Toupee	.69	**	.88	***	.71	**	.69	**	.94	***

Note: * p < .05, ** p < .01, *** p < .001

Within the criteria, the correlation Bald / Toupee ranges from.74 (Attractiveness) to.79 (Self-assurance) to.94 (Health). This means that wearing a toupee changes the relative position only slightly, and it remains almost identical for health. The first two data columns contain a remarkable result: The correlation of the attractiveness ratings A Bald / A Toupee is.74; of the 8 other correlations in which the hair status and/or the criterion are different, 5 are even slightly to notably higher. The smallest correlations are for the cross-ratings S Toupee / H Bald (.53) and S Toupee / A Bald (.56). Only the smallest value fails to reach the significance threshold, but this is irrelevant due to the small number of 13 stimulus persons. More interesting are some findings from Table 2, which we would like to highlight.

The toupee leads to a deterioration in *attractiveness* in only one case, namely number 5. The loss amounts to 0.45 points, which is quite considerable. In terms of natural appearance, he ranks 5th, with a toupee he falls back to 11.5. For the others, the toupee leads to a gain of between 0.21 (3) and a whopping 1.65 points (9). The two men who started best achieved the third and second highest gains with toupees (1.45 and 1.08). The two men at the bottom only gained 0.37 and 0.45 points. On average, the gain is 0.65 points.

As regards *self-assurance*, 5 of the 13 men scored worse with a toupee than with a bald head (3, 4, 13, 7, 10). For No. 3, the loss of 1.12 points is enormous, for No. 10 the loss of 0.03 points is meaningless. The gains are between 0.20 and 0.62 (5, 11). On average, the difference is 0.03 points.

For *health*, 4 men scored worse when wearing a toupee. The loss amounts to 0.53 (3), 0.30 (13), 0.20 (7) and 0.17 points (7). The gains range from 0.05 (12) to 1.07 (9) points. On average, there is a gain of 0.24 points.

If one adds up the gains and losses within the individuals, there are 4 losers: -1.44 (3); -0.43 (13); -0.42 (5) and -0.12 (4). In contrast, there are 5 people who have gained at least one point in total: 3.16 (9); 2.52 (2); 2.41 (1); 1.89 (6) and 1.11 (8). They all recorded a gain in each of the three criteria.

In order to determine the differences between male and female judges, a t-test was carried out for each photograph. In addition to Student's t-test, Welch's t-test and the Mann-Whitney U-test were also run and showed almost perfect agreement. Table 4 shows Cohen's d – not the point difference – for the significant effects ($p < 0.05$). Positive values show higher ratings by the male judges; the reverse is true for negative values.

Table 4. *Effect size of difference between male and female raters.*

	Attractiveness		Self-assurance		Health	
ID	Bald	Toupet	Bald	Toupet	Bald	Toupet
1		0.39		-0.16		
2				-0.21	-0.18	
3	0.31		-0.22			
4		0.55		0.21		
5	0.47	0.61	-0.13			
6				-0.13	-0.17	
7	0.44	0.38				
8	0.55	0.48	-0.17			
9	0.37	0.47				
10	0.52		0.18			
11	0.35	0.43	0.15	-0.14		
12		0.62				0.24
13	0.73		0.17			0.18

In terms of *attractiveness*, 16 of the 26 comparisons show a significant sex difference. In all cases, the ratings by males are more favorable than those by females and it makes no difference whether the men wear their toupee or not. The effect size is considerable, ranging from 0.31 to 0.73; the median is 0.47.

In terms of *self-assurance*, 11 comparisons are significant. In 7 cases, the women gave more favorable evaluations. The effect size ranges from 0.13 to 0.22.

The *health* ratings yield only 4 significant differences. When bald, 2 men are rated higher by the women; the opposite is true for the toupee condition. The effect size ranges from 0.17 to 0.24.

This gives the following overall picture. When it comes to attractiveness, men's and women's views diverge in many cases. Without exception, the ratings by women are (even) more unfavorable than those by men and the effect size is medium to very high. In the case of self-assurance and health, the sex effects are balanced and the effects are weak to medium. For health, despite the very high test power, only a few differences are significant

Current mood of the participants (PANAS)

Following the face assessment, the participants were asked to assess their own current mood using the PANAS. As mentioned, this part of our study serves a different purpose. We would like to briefly sketch the main findings. In an exploratory factor analysis of the 20 items with maximum likelihood extraction and oblimin rotation, two factors emerge in each of the three samples (attractiveness, self-assurance, health) that do not deviate markedly from orthogonality and the correlation shows a negative sign in each case (-0.16; -0.14; -0.10). In each sample, the Positive Affect Scale and the Negative Affect Scale have very high reliability. The mean score for Positive Affect is 2.78, 2.73 and 2.80 (just below the center of the scale, which ranges from 1 to 5). The values for Negative Affect are 1.57, 1.61 and 1.62. As the lower end of

the scale is 1, this means that the test participants experienced virtually no negative feelings at all – which of course would not have been intented. There is no correlation between their current mood and their judgements of the faces.

Discussion

„Toupee or not Toupee?" That is our question. Unlike Hamlet, we are not concerned with existential questions. We are merely interested in the effect of a fully developed androgenetic alopecia compared to a full head of hair. For us, it is not about the perspective of those affected, but about the eye of the beholder when assessing unfamiliar men. Our answers do not imply any recommendations for action for those affected, but this does not exclude the possibility that they may also be of interest to them.

Our focus is on *physical attractiveness* and we had clear expectations in this regard.

Firstly, we expected the men to be seen as more attractive when wearing their toupee. This is definitely the case. The difference is considerable and, statistically speaking, the effect size is strong. We will discuss this in more detail later. Here we want to clarify another point. 12 of the 13 men achieve a gain with a full head of hair, but one appears less attractive with his toupee than in his natural bald state. This man wore his natural fringe of hair at the nape of his neck at an unusual length of around 15 centimeters and the hair was somewhat tousled. In keeping with this, the toupee was also a little tousled and the overall effect was quite eccentric. With trimmed neck hair and a discreet toupee, the effect would certainly not have been negative.

Secondly, we had expected the men to be judged worse by women than by their same-sex counterparts. This assumption was also corroborated. It even applies to every single man, regardless of whether he is seen bald or with a full head of hair. Overall, the effect size is quite high.

Furthermore, we had expected the level of attractiveness ratings to be low. This expectation was also confirmed, but the level is startlingly low. We will come back to this point in more detail.

At this point, we can state: When it comes to*physical attractiveness*, the answer to our question „Toupee or not Toupee?" is a definitive *Sic!*

We had expected a similar picture for *health*, but with weaker effects. This was also confirmed. The men appear healthier with a full head of hair than with a bald head, but the difference is not large. Moreover, it is mainly in the eyes of the men; from a female perspective, the effect size is small. When it comes to *health*, the answer is a lukewarm *Sic*.

We did not specify an expectation for *self-assurance*. At the global level, it does not matter whether the men wear their toupee or not, and at the global level it does not matter whether they are judged by men or women. In terms of *self-assurance*, our answer is *Non*.

The fact that our expectations were confirmed is hardly exciting. Our findings provide further support for the picture that emerges from previous research. Moreover, they are in line with experiences from everyday life, so that even attentive

laypeople would expect something similar. Of course, this does not mean that our study did not provide any interesting insights.

One of the strengths of our study is the explicit consideration of the group level and the individual level. The benefit of this distinction is particularly evident in the assessment of *self-assurance*. At the group level, it does not matter whether the men are bald or have a full head of hair. Nor does it matter whether they are being judged by men or women. At the individual level, on the other hand, we get a differentiated picture. Some target persons are rated significantly better by women and others by men. Interestingly, in 9 out of 10 cases, a sex difference occurs in only one of the two hair conditions. However, it seems unlikely that this pattern can be generalized, especially as the effect size is small. When it comes to *health*, on the whole men differentiate a bit more than women between baldness and a full head of hair. However, the difference is only significant for 4 target persons and the effect is weak. The situation is completely different for *attractiveness*. 11 of the 13 men were rated more negatively by women than by their same-sex counterparts, 5 of them under both hair conditions, and the effect size is medium to very high. Thus, our second hypothesis is not only confirmed at the group level; the extraordinarily harsh assessment of male attractiveness by women affects almost all men. From the individual-centered perspective, there is another point to address. As one would expect, the gain in attractiveness through the toupee is not the same for all men. Interestingly, the two top performers achieved particularly high gains and the two bottom performers only modest gains. Whether this is a peculiarity of our sample or whether it is generalizable along the lines of „Success breeds success" is a fascinating question that, to our knowledge, has not yet been asked, let alone empirically answered.

Looking at the individuals also provides information about the*attractiveness stereotype*, although this is not the focus of this study. When the men wear their toupee, there is a considerable correlation of.69 between attractiveness and self-assurance. When they are seen bald, the correlation is a whopping.87, which is higher than when they are judged on the same criterion under the two hair conditions. For attractiveness, the correlation is.74, for self-assurance.79. For health, the correlation with attractiveness is even closer. All this is all the more remarkable as the traits were assessed by different samples. The more attractive the men are rated by one sample, the more self-assured and healthy they appear to another one. Our study thus provides a further example of the pervasive effect of the attractiveness stereotype. However, it is well known that the motto „What is beautiful is good" cannot be generalized to all personality traits. In our review, we have outlined some findings from studies on male pattern baldness and the literature on the attractiveness stereotype in general fills bookshelves.

Just like the attractiveness stereotype, *age* was not the focus of this study. Nevertheless, it also makes a small contribution to this question.

One aspect concerns the *age of the stimulus persons*. One might assume that older men were particularly devalued, but this is not the case. The order of attractiveness does not coincide with age. For example, the oldest man ranks third in the bald condition and sixth in the toupee condition, while the bottom two are among the youngest. A father ranks one place above his son. The top-ranked is one of the youngest, the second-ranked belongs to the middle group. At this point, it is worth looking at the eldest. As mentioned, with a bald head he achieves 3rd place. The toupee increased his

attractiveness score, but it is only a modest 0.21 points and the second-lowest gain. The change in self-assurance is -1.12 points and in health -0.53. In both cases, this is the biggest loss of all.

The reported details may be interesting, but they do not allow any generalization. Although 13 stimulus persons are more than in some other studies in this area, such a small sample does not permit inferences on the impact of age. A conclusive study would need an entirely different dimension. For example, if one wanted to fill each 5-year interval in the age range from 30 to 70 with 25 stimulus persons, 200 volunteers would be required. 25 people per interval would not be many, as the men also differ in numerous other characteristics, such as hair color, facial hair, eye color, glasses, complexion, head shape, symmetry, configuration and shape of eyes, nose, mouth... As regards the age of the people being assessed, extensive research is still called for.

The same applies to the *age of the judges*. We have unusually large samples, but in this respect they are far from sufficient. In online experiments, the older participants are usually the Achilles' heel. If we wanted to fill every 5-year interval with 25 people in the upper age range as well, we would need tens of thousands of participants with our resources. In order to obtain reliable information on the impact of the age of the judges, one needs access to other populations. This would certainly be worthwhile, because it is safe to assume that the age plays a role that should not be neglected. Unfortunately, there are no conclusive studies available.

In retrospect, the sample size highlights a shortcoming of our study. For self-assurance and health – with regard to our core questions – a smaller sample would have been sufficient and we should have considered additional criteria instead. In the case of attractiveness, on the other hand, a larger sample would have been desirable due to the heavily skewed age distribution. Given the self-selection in the pool we are able to reach, even a sample of 365 is not sufficient for some important questions. However, it is quite adequate for our core questions.

Now we would like to take up a methodological aspect of our core question. We have seen that a toupee increases attractiveness by 0.71 points overall, which corresponds to an impressive effect size of 0.67. Our answer is therefore a clear *Sic!* However, we have also seen that the effect is not the same for all men; for some, the toupee gives an enormous increase in attractiveness, for others the effect is negligible, and in rare exceptional cases a toupee can even have a negative effect. We would like to elaborate on this important issue with the help of Table 5.

Table 5. *Mean scores and ranks. Fictitious scenarios.*

Bald			Toupee			Bald + 0,71	
Rank	Points		Points	Rank		Points	Rank
1	3,35		4,43	1		4,06	1
2	2,95		4,40	1		3,66	1
3	2,87		3,08	3		3,58	1
4	2,83		3,36	1		3,54	1
5	2,62		2,17	9		3,33	2
6	2,57		3,57	1		3,28	2
7	2,37		2,66	5		3,08	2
8	2,32		3,04	2		3,03	2
9	2,28		3,92	1		2,99	2
10	2,17		2,51	7		2,88	3
11	1,84		2,58	6		2,55	7
12	1,73		2,10	11		2,44	7
13	1,72		2,17	10		2,43	7

The first three columns have been taken from Table 2. The first column shows the rank in the natural state of baldness. The second column shows the corresponding score. The third column shows the score when wearing a toupee. The fourth column is obtained as follows: Let's imagine that the man in the respective row joins the group of the others who are all bald and he himself is wearing his toupee. Then he would achieve the rank shown in the fourth column. Not only No. 1 and No. 2 would then be at the top, but also No. 4, 6 and 9. No. 8 would be in second place and so on. But No. 12 and No. 13 would only come in 11th and 10th place. No. 5 would even drop from 5th to 9th place. The two columns on the right are based on the following consideration. The average gain due to the toupee is 0.71 points. Now we imagine that each bald man would achieve exactly this gain. This would yield the rankings in the right-hand column. In this case, the relative improvement would be fantastic for most and a considerable benefit even for the bottom two. The comparison with the preceding scenario makes it clear once again that not everyone benefits to the same degree from a toupee and that the men at the bottom in particular have a smaller gain. Considerations of this kind would certainly be more valuable for those affected than statistical significance and effect sizes.

At this point we suggest another approach to the interpretation of effect sizes, which should be easy to understand, at least for psychologists. The intelligence quotient IQ has a standard deviation of 15. Cohen's d indicates the mean difference in standard deviation units. An effect size of d = 0.71 corresponds to a difference of 10.7 IQ points – that's a huge difference! At the level of nations, for example, this corresponds to the difference between Switzerland, Great Britain and Germany on the one hand and Turkey, Thailand and the United Arab Emirates on the other. Nobody would mistake these two groups of countries for each other, neither in terms of national intelligence nor in terms of countless other characteristics that are to a considerable extent a consequence and, to a certain extent, a cause of the differences in intelligence. An overview of national IQs based on the current state of research can be found in Henss (2023).

We have saved one very important topic until here, namely the extremely low attractiveness ratings. At the group level,

the mean of the bald men is 2.39. The means for the individuals range from 1.72 to 3.35. Here we should remember that the grades of the rating scale were, as usual, coded with values from 1 to 9. As this is a bipolar scale with the poles „unattractive“ and „attractive“, it would be better to imagine a scale from -4 to +4. Then the means of the individuals would range from -3.28 to -1.65 and the overall mean would be -2.61. The least attractive man is only 0.72 points above the lower end of the scale and even the most attractive one is far from the neutral center. Even with his toupee, the most attractive man would still have received a negative score (-0.57). If one were to take these ratings at face value, one would come to the conclusion that the men were all unattractive, some even extremely unattractive and ugly. Such a conclusion would in no way be justified.

It should be emphasized once again that our stimulus persons are quite ordinary men, as one encounters them in public in everyday life. None of them are conspicuously unattractive and none are conspicuously attractive. The men were also by no means ungroomed. On the contrary, they attached great importance to their appearance and wore an expensive toupee to match their type. The small section of clothing visible in the photograph made a well-groomed impression. According to our assessment, the men would have received scores from 3 to 7, or in the more appropriate perspective from -2 to +2. To differentiate as strongly as possible, we might even have awarded an 8 in one or two cases and maybe a 2 in one case, but definitely neither a 1 nor a 9.

The enormous difference between our own standards and the assessment by hundreds of raters cannot be attributed to a „cautious“ response tendency on the part of the raters and certainly not to an inappropriate positive bias on our part. The judgments of self-assurance and health are, as one would expect for a sample of quite average stimulus persons, perfectly in the middle of the scale. It is precisely this discrepancy that makes the devaluation of men’s attractiveness crystal clear. Men are not generally rated negatively, but they are denied attractiveness. This applies not only to men who have lost their hair, but also to men with a full head of hair, and it holds all the more so in the eyes of female judges.

This highly interesting phenomenon is not specific to this study. We have conducted numerous experiments on face perception with different sets of photographs (e.g. Henss 1992, 1998). Most of them online with hundreds or thousands of participants and often in a German and an English version. In most cases when photographs of men were taken into account, the devaluating assessment of their attractiveness was noticeable, although not quite as extreme as in this study. The attractiveness malus for men is not a curiosity of our own investigations; it can be found in large parts of the literature on attractiveness research, but is strangely not addressed at all or only in passing.[11] If there were an attractiveness malus for women, there would be a roaring, ear-splitting outcry. The derogatory judgment of male attractiveness is usually not even alluded to.

I would like to conclude with a personal remark. As mentioned, almost a quarter of a century ago I outlined the state of research into the social perception of male pattern baldness up to the year 2000 in a review (Henss, 2001). In it, I lamented the paucity of research and pointed out numerous weaknesses. Since then I have dealt with other topics and I have not currently carried out a systematic literature search. To all appearances, the situation has not improved; on the contrary, it seems to have worsened. Recent empirical research seems to be scarce and most of the few studies I know of have minor or even serious shortcomings. Overall, the field of social perception of hereditary hair loss in men seems to

me to be in a shockingly poor state. I think men deserve better than that.

Statements and Declarations

The data sets in omv format are freely available at https://www.researchgate.net/profile/Ronald-Henss

Other formats are available on request.

Footnotes

[1] https://www.researchgate.net/publication/244915177_Social_Perceptions_of_Male_Pattern_Baldness_A_Review

[2] „[E]vidence, which speaks primarily to natural male pattern baldness, has little to say about men who choose baldness" (Mannes, 2012, p. 3) and vice versa.

[3] On the history of the battle against hair loss, see for example Campo and D'Acunzo (2016), Homan (2019), Kligman and Freeman (1988), Seagrave (1996). One brave fighter should be highlighted in particular, namely Julius Caesar. „First, he grew his thinning mane long in the back and brushed it over his scalp, in an early version of the comb-over. When that did not work (hairspray had yet to be invented), his lover Cleopatra recommended a home remedy consisting of ground-up mice, horse teeth and bear grease. This too had little effect, so the Roman dictator took to covering his scalp with a laurel wreath" (Homan, 2019, S. 25).

[4] On the broad spectrum of modern methods, some of which have not yet entered public awareness, see for example Asfour, Cranwell and Sinclair (2023), Kelly, Blanco and Tosti (2016), Nestor et al. (2021), Qi and Garza (2014).

[5] On the medicalization of male hair loss, see for example Jankowski and Frith (2022).

[6] On the diverse effects on social perception, see Henss (2001).

[7] On the psychological impact of hair loss on those affected, see for example Alfonso et al. (2005), Budd et al. (2000), Cash (2009), Frith and Jankowski (2023).

[8] „We applied and further adapted black-and-white portrait pictures that had been created by Neave and Shields (2008) with a facial composite software. In particular, the original 10 nonbald targets (Caucasian men in their twenties with full short hair of different color and texture) were additionally morphed into complete bald targets" (Kranz et al., 2019, p. 3).

[9] The photographs were provided by the Swiss Parliamentary Service.

[10] The German word *attraktiv* usually refers to the physical aspect, so it should be translated as physically attractive, handsome or good-looking. The opposing pole to *selbstsicher* was *unsicher* (insecure). It should therefore be translated as self-assured or self-confident, not self-conscious, which has a different connotation. We will use the term self-assured throughout.

[11] In our work, we have explicitly pointed out the sex difference and spoken of an attractiveness bonus for (young) women (in particular Henss, 1992; Henss, 1998). Since the women's scores are rarely much above the middle of the scale, but the men's scores are usually more or less clearly below it, it seems more appropriate to speak of an attractiveness malus for men. Furthermore, we have repeatedly pointed out an interaction between the sex of the judges and the sex of the judged, which we have termed own-group bonus. In terms of attractiveness, one's own sex is generally judged relatively, if not absolutely, better than the other sex. As only men were assessed in this study, it cannot contribute to this effect. Hundreds of empirical studies are available from the most diverse areas of face perception. It would be a worthwhile task to investigate both effects as part of a meta-analysis.

References

- Alfonso, M., Richter-Appelt, H., Tosti, A., Sanchez Viera, M. & García, M. (2005). The psychosocial impact of hair loss among men: a multinational European study. *Current Medical Research and Opinion*, 21:11, 1829-1836. https://doi.org/10.1185/030079905X61820
- Asfour, L., Cranwell, W. & Sinclair, R. (2023). Male androgenetic alopecia. *National Library of Medicine. National Center for Biotechnology Information*. https://www.ncbi.nlm.nih.gov/books/NBK278957/ Accessed 23.03.2024.
- Budd, D., Himmelberger, D., Rhodes, T., Cash, T.E. & Girman, C.J. (2000). The effects of hair loss in European men: a survey in four countries. *European Journal of Dermatology*, 10 (2), 122-7.
- Buss, D. (2016[4]). *The evolution of desire: Strategies of human mating* New York: Basic Books.
- Campo, D. & D'Acunzo, V. (2016). Doctors and baldness: a five thousand year old challenge.*Giornale Italiano di Dermatologia e Venereologia*, 151 (1), 93-101.
- Cash, T. (2009). Attitudes, behaviors, and expectations of men seeking medical treatment for male pattern hair loss: Results of a multinational survey. *Current Medical Research and Opinion*, 25 (7), 1811-1820 https://doi.org/10.1185/03007990903005201
- Chan, H.F., Skali, A., Stadelmann, D., Torgler, B. & Whyte, S. (2020). Masculinity cues, perceptions of politician attributes, and political behavior. *Economics & Politics*, 2020; 00, 1-24. https://doi.org/10.1111/ecpo.12164
- Dion, K.K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, 24, 285-290. https://doi.org/10.1037/h0033731
- Eagly, A.H., Ashmore, R.D., Makhijani, M.G. & Longo, L.C. (1991). What is beautiful is good, but..: A meta-analytic review of the physical attractiveness stereotype. *Psychological Bulletin*, 110, 109-128. https://doi.org/10.1037/0033-2909.110.1.109
- Frith, H. & Jankowski, G.S. (2023). Psychosocial impact of androgenetic alopecia on men: A systematic review and meta-analysis, *Psychology, Health & Medicine*. https://doi.org/10.1080/13548506.2023.2242049
- Henne, S.K., Nöthen, M.M. & Heilmann-Heimbach, S. (2023). Male-pattern hair loss: Comprehensive identification of the associated genes as a basis for understanding pathophysiology. *medizinische genetic,* 35 (1), 3-14. https://doi.org/10.1515/medgen-2023-2003

- Henss, R. (1991). Perceiving age and attractiveness in facial photographs. *Journal of Applied Social Psychology*, 21, 933-946. https://doi.org/10.1111/j.1559-1816.1991.tb00451.x
- Henss, R. (1992). *„Spieglein, Spieglein an der Wand..." Geschlecht, Alter und physische Attraktivität* Weinheim: Psychologie Verlags Union.
- Henss, R. (1997). The Big Five and beyond. Poster presented at the *VIII. Biennial Meeting of the International Society for the Study of Individual Differences*, Aarhus, Danmark, 19.-23. July 1997. https://www.researchgate.net/publication/337312566_The_Big_Five_and_beyond
- Henss, R. (1998). *Gesicht und Persönlichkeitseindruck*. Göttingen: Hogrefe. [Reprint: Henss, R. (2020). *Gesicht und Persönlichkeitseindruck*. Saarbrücken: Dr. Ronald Henss Verlag. Also freely available as PDF https://www.researchgate.net/publication/339147401_Gesicht_und_Personlichkeitseindruck
- Henss, R. (2001). Social perceptions of male pattern baldness. A review. *Dermatology and Psychosomatics / Dermatologie und Psychosomatik*, 2 (2), 63-71. https://doi.org/10.1159/000049641 Also https://www.researchgate.net/publication/244915177_Social_Perceptions_of_Male_Pattern_Baldness_A_Review
- Henss, R. (2023). The Intelligence of Nations. National IQs. Update 2023. *Qeios*. https://doi.org/10.32388/6LBP4T
- Jankowski, G.S. & Frith, H. (2022). Psychology's medicalization of male baldness. *Journal of Health Psychology*, 27 (9), 2161-2180.
- Kelly, Y., Blanco, A. & Tosti, A. (2016). Androgenetic alopecia: An update of treatment options. *Drugs*, https://doi.org/10.1007/s40265-016-0629-5
- Kligman, A.M. & Freeman, B. (1988). History of baldness from magic to medicine. *Clinics in Dermatology,* 6 (4), 83-88.
- Kranz, D., Nadarevic, L. & Erdfelder, E. (2019). Bald and Bad? Experimental evidence for a dual-process account of baldness stereotyping. *Experimental Psychology* (2019). https://doi.org/10.1027/1618-3169/a000457
- Mannes, A. (2012). Shorn scalps and perceptions of male dominance. *Social Psychological and Personality Science*. 4, 198-205. https://doi.org/10.1177/1948550612449490 https://xgayru.info/misc/docs/0/289_443517.pdf Accessed 23.03.2024.
- Neave, N. & Shields, K. (2008). The effects of facial hair manipulation on female perceptions of attractiveness, masculinity, and dominance in male faces. *Personality and Individual Differences*, 45 (5), 373-377. https://doi.org/10.1016/j.paid.2008.05.007
- Nestor, M.S., Ablon, G., Gade, A., Han, H. & Fischer, D.L. (2021). Treatment options for androgenetic alopecia: Efficacy, side effects, compliance, financial considerations, and ethics. *Journal of Cosmetic Dermatolology*, 2021; 00:1-23. https://doi.org/10.1111/jocd.14537
- Qi, J. & Garza, L.A. (2014). An overview of alopecias. *Cold Spring Harbor Perspectives in Medicine*, 4 (3). https://doi.org/10.1101/cshperspect.a013615
- Seagrave, K. (1996). *Baldness: A social history*. Jefferson, NC: McFarland & Co. Inc.
- Wade, J.T.,·Fisher, M.L. &·Burch, R.L. (2021). Toupee or not toupee?: Cranial hair and perceptions of men's attractiveness, personality, and other evolutionary relevant traits. *Evolutionary Psychological Science*. https://doi.org/10.1007/s40806-021-00303-y