ON THE PSYCHOLOGY OF JUDGING CATS

Peter Paul Moormann

Introduction

Subjectivity is a core element in all areas of human judgement and cat exhibitions are not immune. Judging pedigree cats is a particularly interesting field for studying the effects of subjectivity in human judgement, whereby breeders, club members, judges and spectators all may display differences of opinion on which cat deserves a title, Best of Breed (BOB) or Best in Show (BIS) and why. Much debate and speculation takes place as to why certain awards are given or withheld. Sometimes it is assumed that a particular judge has awarded or withheld a title deliberately for reasons other than those having to do with the quality of the cat. The judge is simply not honest. For example for reasons of favouring cats descending from the judge's own breeding lines, liking or disliking certain cat owners or cat breeds, matters of reputation, soliciting for being invited by the president of a prominent club, etc. In those cases we are talking about cat politics. However, sometimes judges make decisions which lead to disapproval by other cat experts (e.g. leading breeders and other judges), but which are not made deliberately.

1. This article will focus on the different kinds and causes of such unintentional, often unconscious biases and the settings in which they are likely to occur.
2. Furthermore the three most influential judging systems (e.g. GCCF system from Great Britain, the traditional system from the continent of Europe, and the ring system from America) will be discussed. Although being different in the way subjectivity is dealt with, it should be noted that the three systems have in common that all are meant to harmonise individual bias involved in cat judging.
3. Then an alternative, more objective method will be presented for choosing BIS.
4. Finally it is concluded that pointing out the existence of involuntary biases and the psychological mechanisms underlying them is the first step to be made in an attempt to prevent cat judges from making such errors.

The process of judging cats

Judging cats is a beauty contest. How the ideal cat should look like is written in the *Standard of Points*. For each breed descriptions are given on several dimensions, in relation to the phenotype of the cat:

1. Form (body, head, ears, legs, feet, tail, eye shape), which defines the type of the cat.
2. Coat colour, colour distribution (torties, bi-colours), and pattern (tabby, ticking, tipping).
3. Coat texture (silky, woolly, dense, curly, naked, long, short, undercoat, etc.).

4. Eye colour

5. Temperament is mentioned in a few cases. For instance in the GCCF Standard of Points for the Ragdoll: “The cat should be relaxed in temperament, gentle and easy to handle”.

In most cases temperament should be judged according to what is written in the show rules. Bad tempered cats generally are disqualified. The same holds for inadequate grooming and the presence of fleas and other parasites. Furthermore in the GCCF for instance, a standard list of withholding faults, concerning all breeds, is given. In kitten open classes certificates or first prizes ought to be withheld for any exhibit, which is:

1. Undersized for its age and breed.
2. Not in excellent physical condition.
3. Not free from physical defects as defined in the list of defects* (entropion, monorchism, kink, narrow or twisted lower jaw, etc.).

*Folded ears are mentioned as one of the defects. Therefore Scottish Fold and American Fold are not recognised by the Governing Council of the Cat Fancy (GCCF).

For each breed the relative importance of the subdimensions (head, body, tail, etc.) within each dimension is expressed in points. Hence from a judge it is expected that he/she gives points to all the subdimensions, computes the sum total of all the dimensions involved and decides according to the height of the sum total whether the title is awarded or withheld. Furthermore when comparing several cats in a class the sum total of each cat determines its rank order. This is all what is asked for in the standard of points system. Although some judges actually do give points the majority of the judges consider the points attached to each subdimension of the standard more as an indication for the relative importance of that specific subdimension in the whole picture of the cat. Nevertheless *the standard of points system implies an analytical assessment whereby all the sudimensions of one particular cat are simultaneously compared with all the subdimensions of all the other cats within a group of cats.* This is a highly complex mental operation and *it could be questioned whether normal human beings possess the required abilities to carry out such a difficult task.* Is our memory capacity equipped well enough to fulfil the demands of this gigantic mental enterprise?
In the psychological literature contradictory findings can be observed on a person's ability to evaluate so many dimensions simultaneously. Stockholders are inclined to use 6 to 7 dimensions of the total information, radiologists use 2 to 6 dimensions, and lawyers only use 1 to 3 dimensions. However, in all those cases much more information was available then was actually used. *What all judges had in common was that they did not take all the relevant dimensions equally into consideration.* The expert judge appears to select and so use only a part of the information available. This might explain why some cat judges, are inclined to judge according to type of head alone. Adding colour and coat texture makes the tasks more complex. Individual differences in cognitive complexity might play a role as well. Cognitively complex individuals are more capable of assessing many dimensions at the same time than cognitively simple individuals. Moreover, results

from psychological studies show that experts in judging combined 9 to 11 dimensions in a laboratory situation, but not more than 3 in a natural situation. As cat shows happen in natural situations it seems likely to assume that *the average judge cannot handle more than 3 dimensions simultaneously*. The three most important ones are type, colour and coat texture. With these three main dimensions in mind a judge must be able to differentiate both between the different breeds in case of BIS and within colour classes of one breed in case of BOB.

Are there any scientific studies published on cat judging? None as far as I know. From observation and introspection I have noticed two different styles in cat judging:

1. The analytic style.
2. The holistic style.

The philosophy behind the analytic style is based upon the idea that *the whole is the sum of its parts.* In fact this is exactly what is stated in the Standard of Points. Points are assigned to each part of the cat. The sum total of all the different parts is used as the final result. It is an additive, highly rational, and analytical model, which primarily relies on left hemispheric activity, responsible for language and calculating skills, where information is processed sequentially (such as in reading, writing and computing). A judge with an analytic style tries to express the qualities of the cat in points. The analytic judge is convinced that numbers (e.g. points) are more objective, but forgets that there is no objective measuring device available other than his/her own brain. Hence it remains subjective. Assigning points to the cat is a quasi-objective activity. Sometimes the analytic judge may be surprised by the result of his/her own quasi-objective qualifications. When looking at the judging sheet cat Nr. 1 got more points than cat Nr. 5, and should therefore win. However when both cats were presented simultaneously to the judge cat Nr. 5 looked better than cat Nr. 1. What to do now? Leave it that way or change his or her decision. Can the result be attributed to a miscalculation?
Not necessarily, because *the whole can be more than the sum of its parts*, e.g. the philosophy behind the holistic style. The holistic or Gestalt style assumes that judges are inclined to start with a global impression of the cat. Is this a good, a very good, or an excellent cat? This global impression is based upon a match between the ideal mental image of that particular breed and the actual appearance of the cat in question. The closer the match fits the ideal mental image the higher the appreciation of the judge will be. It is interesting to note that global impressions rely on right hemispheric activity, the hemisphere involved in visuo-spatial orientation, intuition and parallel processing of information (such as in looking at a painting in an art gallery). However a holistic judge may experience some problems as well: “Chinchilla Nr. 21 looks better than Nr. 22. It has a better type and beautiful light tipping”. However, a closer, more detailed look reveals a problem. Nr. 21 has yellow eyes, is lacking the maquillage and brick red nose, which is the hallmark of a good chinchilla. Nr. 22 however, has a super bluish-green eye colour and a nice open face despite the somewhat longer nose. “Well, I have to change my mind. According to the standard Nr. 22 is better than Nr. 21”.
The point to be made is that both styles can be involved in judging. Although a judge may have a preference for one style *most individuals are flexible enough to switch from*

one style to the other. Why? Because judging styles are assumed to be associated with hemispheric specialisation, and because communication between the left and the right hemisphere is an essential feature of the normally functioning brain. The transmission of information from left to right and vice versa enables the use of both the analytic and holistic style. It is nothing special. The right hemisphere is responsible for an instant global impression and the left hemisphere is involved in a more detailed analysis of the cat. Intuitive processes are handled in the right hemisphere and may lead to the feeling: "This is the cat I have to choose". The left hemisphere helps us in finding a rationale for our decisions and protects us against impulsive acts not based on logic. Furthermore both hemispheres communicate with each other through the corpus callosum. This brain structure enables us to find words for non-verbal material, such as the ideal visual image of the Siberian cat, or to construct a visual image of written material derived from cat standards we are not familiar with. Hence the analytic style is predominantly associated with left hemispheric activity and the holistic style with right hemispheric activity. As both hemispheres are able to communicate which each other in a variety of situations there is no reason to suppose that they would not communicate with each other when judging cats. However, a judge who primarily relies on the analytic style focuses too much on *details* (for instance preoccupied with defects) and is apt to miss the balance in a cat. A judge who predominantly relies on the holistic style may neglect details (for instance physical defects) and can be become a victim of the *halo-effect*, e.g. the tendency to rely on a more global impression when judging instead of analysing all the relevant dimensions in detail.

Subjectivity

The judging of cats resembles other assessment procedures in not having a more objective measuring device for assessing performance (here judging cats) than the human eye and brain. Although the human brain is highly ingenious machinery, which is in some respects still better equipped than the most advanced computer for carrying out a diversity of mental operations, it has the peculiarity of processing information in an *idiosyncratic* way. This inevitably means among other things that *there is no guarantee that, two or more individuals perceive what happens in the outside world, in an identical way*. This phenomenon is in general referred to as 'subjectivity'.

Perception

Perception is guided by a variety of factors, such as:

a. Judging <u>experience.</u> Individuals with experience in a particular field appear to perceive a situation in a more differentiated way. They simply see more. This allows for finer and better discriminations when judgements are made. A professional wine taster, a 'nose', for instance, knows subtle differences between all kinds of wine. Therefore he is better equipped to perceive and judge the quality of a wine, determine when it was bottled, which grapes were selected, and from which vineyards it originated. The same holds for the judging of cats. *Experience leads to differentiation in perception and better discrimination, which facilitates decision-making.* But what to do when necessary experience is lacking? New breeds regularly appear on shows

and it happens frequently that judges are asked to assess such 'nouveautés'. The judge does not yet have the ideal mental image in mind of the new comer. The only information available is a preliminary standard, in rare cases a drawing or some photographs. Not being familiar with breeds which are new or seldom seen on shows may lead to situations where a judge has to construct his or her own mental image derived from the written standard of that breed without the knowledge usually obtained from the observation of excellent exhibits in vivo. This poorly constructed mental image will direct his/her perception and as a consequence the way the cat is judged. The worst outcome is a completely different interpretation of the standard, or judging according to norms not shared by the inventors of the breed. However, highly experienced all-breed judges are more likely to do a better job. The can rely on the knowledge of features characteristic of breeds close to the new breed. What is the essential difference between the new breed and already existing breeds? Where is overlap and where are the differences? *In fact the all-breed judge has a frame of cat types in mind, a theory on how lines can be combined into forms that define different cat types.* This Expert's Knowledge enables him/her to find a proper location for the new comer in the already existing picture gallery of cat images the judge has in mind. The mental picture gallery of all-breed judges is comparable with the art collection of a famous museum, where many different schools of painting are represented: Dutch 17th century paintings, impressionism, expressionism, cubism, surrealism, etc. The mental picture gallery of judges with a restricted package is comparable with the collection of a local museum in the countryside, where only one school is represented. *From the above it is concluded that only all-breed judges should be invited for Best Overall Cat-panels (short, semi-long, and long).* They possess more experience and a more elaborated Expert's Knowledge. Therefore they are better equipped to perceive the subtle variations needed to discriminate between the different breeds, and to judge each exhibit on its general qualities regardless of breed.

b. Personality traits. *Judging cats is assessing beauty.* Harrison Weir (1824-1905) was the first person that tried to measure beauty by introducing his 'Points of Excellence', the standards for judging, at the first cat show held at Crystal Palace in London on 13 July 1871. Harrison Weir, the Father of the Cat Fancy, was a painter, a cat illustrator. Furthermore he was a judge at poultry and pigeon shows. From Harrison Weir it can be said that he had an 'eye' for beauty, a sense for aesthetic values, a feeling for lines, forms, colours, and proportions. However *certain personality traits affect the way we look at or perceive beauty*. Individuals with a rigid personality structure stick to the old. They feel safe with what is old, what is familiar. It gives them a sense of trust. They are afraid of new developments, for they have the feeling that it may lead to loss of control. However what defines a top Persian or a top Siamese, e.g. *the prototype of the ideal Persian or Siamese is not a static construct. It is dynamic, develops and changes over time*. Judges with a *rigid* personality are less willing to accept those changes, not prepared to alter their perception. They stick to more *traditional*, old-fashioned ideas of an ideal cat race, which may hamper the progression toward the perfect specimen of a particular breed. The opposite of rigidity is *flexibility*, or openness to new experiences. However some individuals don't know where to slow down in their pursuit for new experiences. They cannot stop and have to go on. Judges with a weak stop system in their personality make-up have a tendency towards

overshooting. They go for the *extreme, overtyped* cats. A tendency, which eventually will end up in only two types, e.g. The Persian at the one end and the Oriental at the other end of the scale. A form of globalism of cat types. We get a lot of the same and there are not many nuances in between. The *globalism of types* is no fiction but already reality. The Turkish Angora and Norwegian Forest Cat are becoming more Oriental in appearance and the British, Burmese, Burmilla, and Asian are becoming more like Exotics or Persians (particularly in the USA). Therefore attitudes, being too conventional on the one hand and being too renewing on the other hand should be controlled. The car should be kept in the middle of the road. Driving too much to either the right or the left will inevitably lead to a crash. *Cat judging is a matter of balance.*

c. Physical state. Illness, age, intoxication, and tiredness all affect perception, concentration, and speed of judging (information processing is slower). This hardly needs an explanation.
d. Mental condition, emotional state and mood are associated with physiological arousal, which in turn influences attention and perception. Extreme *nervousness* or anxiety induces arousal increases that may lead to *involuntary narrowing of attention* (tunnel vision). *Important aspects of the judgement task may therefore not be perceived and overlooked.* Furthermore it should be noted that some personality traits are associated with emotional state and mood. *Neurotic* individuals for instance are more anxious, depressed, have a tendency to worry and experience *difficulties with decision making*. These characteristics not only affect the way the cat is seen, but is *slows down the speed of judging* as well. Another interesting factor is mood. *Mood is likely to affect consistency in judging.* Irritation, anger, and hostility affect perception and appreciation and can lead to an underestimation of an exhibit. Being extremely happy or euphoric, only seeing the bright side of life, easily leads to an overestimation. This phenomenon is called *leniency-severity*, the tendency to assess objects either too high or too low.

Appreciation

Appreciation is not only dependent upon what the judge observes or perceives. Knowledge of an *implicit norm* on what is a good, a very good, or an excellent exhibit is necessary as well. However, some judges are either overcritical or too mild, regardless of mood. This personality characteristic affects appreciation, hampers the consensus among judges, and is called *high versus low judges*. Another factor that influences appreciation is the *dynamic character of norms*. As a breed develops, the norm changes accordingly. This can easily be verified when looking at pictures of Persian cats. Although the written standard has not changed much over time the type of the ideal Persian, for instance, has undergone some remarkable changes. The whole cat has become cobbier and the face has become flatter with a much shorter nose. Comparable developments can be noticed in the Siamese, but in the opposite direction. The contemporary Siamese is finer boned, has a more pointed head with a much longer nose and larger ears than the old type (nowadays the old type is called Thai). Hence, *judging regularly and recently is needed to keep in touch with that up-to-date norm.*

For some breeds it can be argued that they are not yet ready. A good example is the Siberian. The lack of consensus on how the ideal Siberian should look like is expressed in the number of different standards used. What is the prototype of the ideal Siberian? It could be said that the Siberian does not yet have a signature, an identity. Moreover it is hard to distinguish the Neva Masquerade from a Colour pointed Ragdoll or a Colour pointed Persian with a weak type. In my opinion it would have been better if Siamese points were abolished altogether for Siberians. Coloured points lead to confusion, particularly with Ragdolls. Eventually the ideal type of the Siberian will evolve, but only if serious attempts have been made to define the features, which distinguish a true Siberian from all the other semi-longhairs. This was successfully done for the Maine Coon and the Norwegians. In the beginning of the seventies Maine Coons, Norwegians, and Persians lacking type often could hardly be distinguished from each other. They all looked similar. Semi-longhairs with a medium sized wedge. Ordinary household pets. Then breeders and judges refined the standard. Differences in coat texture and type (head shape, profile, ear and eye setting) were emphasised. The results are obvious. Now, after 25 years of selective breeding, one can instantly see what is a Norwegian and what is a Maine Coon. The same ought to be done for the Siberian. Only then judges will appreciate this breed and propose it for BIS. The foregoing clarifies that *judges have a certain power to determine in which direction a cat race is going to develop*. Judges are in a position in which the aesthetic appearance of a breed can be shaped. In addition to power *judges have responsibilities* as well. They not only act as guardians for the aesthetic looks of a cat breed, but they should take care of the health of cat breeds as well. Therefore *physical deformations ought to be penalised.*

Furthermore *appreciation is affected by how unique, well-typed, well coloured, well balanced, well groomed, or charismatic a cat is assessed by that particular judge.* I'll give an example. Eye colour in Colour Points (CPs) has deteriorated in the past decades. Nowadays they are very pale and often have yellowish green outer rims. They old CPs were breeded by crossing Persians with Siamese. The first CPs had weak types (long bodies, thin legs, long tails, wedge-shaped heads and large pricked ears), but a quite good eye colour (not the China blue from the Siamese, but aquamarine blue). Then top blue Persians were used to improve the type. This endeavour has been a partly successful enterprise. Although type and coat texture was improved considerably eye colour was lost, and the points often were ill defined and brindled. Now CPs with excellent types is more or less the rule. However excellent eye colour has become very rare, and SPs with a pale body colour and dark seal points, including the tail, is hardly ever seen. In fact the judges are to blame for the loss of eye colour. They only looked at type and gave titles and BIS for CPs with watery eye colours and yellow in the eyes. Hence the judges encouraged the breeders to concentrate on type at the cost of eye colour. However, a judge who knows to appreciate how difficult it is to get deep blue eyes is prepared to prefer depth of eye colour above type when choosing between two exceptionally good cats. The judge in question might be criticised by naive breeders, and might be accused of disliking typed cats, or being old-fashioned. Pure nonsense, the judge already was one step ahead, and was signalling that CP-breeders should work on eye colour from now on. Such nuances in appreciation imply that the *weights attached to each of the observed elements (type, eye colour, points, etc.), described in the standard of points of that specific breed, and their added scores, may differ for each judge*.

Involuntary bias when judging big classes

When human judgement is involved in assessing classes of cats the judge is usually presented with a whole series of exhibits, one after the other. After each cat the judge is given a limited amount of time to pronounce a judgement. Thereby the judge is expected to display a *consistency* of judgement, so that the better cat is always awarded higher than an inferior cat. However, *when relatively large numbers of cats compete in a class there is a corresponding increase in the information that has to be compared, and a corresponding increase in the probability of giving inaccurate ratings.* The main concern of this paragraph has to do with those systematic errors or involuntary biases involved in judging series of cats. The predecessor effect is one of them, i.e. the tendency to give a qualification that is unconsciously influenced by the qualification given in the preceding judgement. Two variants of the predecessor effect can be formulated:

1. The contrast effect, where the qualification of the foregoing cat has a reinforcing effect on the qualification of the next cat. A very good cat directly following a good cat will be marked relatively too high (excellent), while a very good cat directly following an excellent cat will be marked relatively too low (good). In judgements of physical attractiveness in humans empirical evidence has been brought forward for the latter variant of the contrast effect.
2. The assimilation effect, where the qualification of the foregoing cat has a neutralising effect on the qualification of the next cat. An excellent cat directly following a good cat will be marked relatively too low (very good), while a very good cat directly following an excellent cat will be marked relatively too high (excellent).

Both variants of the predecessor effect should be taken into consideration when judging big classes or groups of cats. *Becoming aware of its existence is the first step to be made when trying to prevent its impact.*

The specific judgement situation in judging big cat classes leads to the conclusion that it can be characterised by the following features:

1. Each cat is judged at a specific point in time.
2. Competing cats are judged in sequence, one after the other.
3. After the first cat, the judge categorises this exhibit based on both an implicit norm regarding the quality of that cat (good, very good, excellent) and the expectations of those cats, which will come later in the series.
4. As the series of exhibits progresses each cat has therefore to be compared with the first cat (anchor), and the other preceding competitors.
5. At the end of the class or group the judge can have a glance at all cats simultaneously either by stewards who are presenting the cats (traditional system) or by having a look at the cats who are placed in a row of show-pens (GCCF and ring system).

The ability to compare an observation with earlier ones and their respective qualifications is primarily a memory task, and therefore dependent upon:

1. Short-term memory (STM) capacity,
2. Long-term memory (LTM) storage, and
3. Long-term memory recall and recognition.

In the comparison of observations many errors may occur. Our short-term capacity does not exceed 7 plus/minus 2 digits. This rather limited capacity is certainly a potential source for inaccurate judgements, simply because *not all information can be processed into LTM.* STM serves as a strong filter. It implies that some aspects of the cat cannot be stored in LTM (for that reason green rims, grey undercoats, unbroken necklaces, lack of contrast in smokes, or white medallions may easily be 'overlooked'). Brief written notes can be helpful to compensate for this memory inconvenience. Studies on LTM-storage and LTM-recall show that *both encoding and retrieval problems, can be caused by interference of similar material.* It can be argued that interference is more the rule than the exception when judging cats in big open classes (GCCF and continental system). Fifteen British Blues for instance look very similar. An untrained eye already experiences difficulties in distinguishing different breeds from each other: for them British Blue, blue European Shorthair, blue Burmese, Russian Blue or Korat all look the same: they are blue and the have a short coat. I even don't dare to mention the Chartreux. However, even a trained eye may experience problems when discriminating between many exhibits belonging to the same breed, sex, and colour is asked for. Quite understandable: the cats have much in common. They don't differ from each other in the kind of stimulus material (they are all British Blue) but instead to the extent that each aspect of the stimulus material (e.g. type of head, coat colour, coat texture, etc.) differs from cat to cat. From a psychological point of view this is an extremely complicated task with a heavy memory load. Therefore judges are inclined to split up the whole group of exhibits into smaller subgroups according to overall quality (good, very good and excellent). As a consequence the demands on memory load become much smaller within subgroups, and interference effects can therefore be attacked more successfully.

Judging systems

We have seen that all kinds of intentional and unintentional biases can be distinguished in the process of judging cats. Judges differ in their ability to perceive, appreciate, and compare observations. The situation in which the cats are judged is a factor of importance as well. Since the beginning of the last century, *numerous judging systems have been developed in an attempt to harmonise individual bias involved in cat judging*. Three systems have survived: the GCCF, the continental, and the ring system. Now a comparison between the three systems will be made.

a. In the British or GCCF system some judges first have a global look at all cats to get an impression of the overall quality of the cats. Then all the exhibits are taken out of the show-pen by a steward and are judged individually on the trolley table. The 4 best exhibits (the number is optional) are selected and judged more precisely. Brief notes

are made on each selected cat and a rank order is given. Then it is decided whether the champion certificate is awarded or withheld. A strong point of the British system concerns *the almost total abolishment of factors that might influence a judge*. The owners are kept away from the pens when the open classes are judged. The cats are anonymous. No prize-winning certificates are allowed in the pens during the judging of open classes. The steward is the only factor, which can influence the judge directly. Moreover it is a *fast judging system*. Written reports are only needed for the best cats. However, not the judging per se, but *the written feedback on the results is a matter of concern for memory distortions*. The reports are written after the show, and are therefore dependent upon what is stored in memory and what is written in the brief notes. From a psychological point of view this is the weakest part of the British system. The delay, the limited amount of written information, and the similarity between the cats are potential sources for memory recall and recognition problems. The judge has to construct a written description derived from brief notes and an already fading mental image of that particular exhibit. Furthermore interference with similar exhibits is likely to occur. This is asking for problems. Fortunately it has nothing to do with the actual judging. The awards have been given long before the writing commences.

b. In the traditional or <u>continental system</u> each cat in the class is judged separately and extensive reports are written for each cat. *The written feedback given to the owners therefore is a strong point of this system.* However it goes *at the expense of the judging speed*. In the past stewards brought the cats to the judging room. That was for reasons of anonymity. Nowadays the owner acts as steward, and it cannot be denied that the *owner's effect is as a serious source for both intentional and unintentional biases. It affects the reliability and validity of judgements. In fact this is a highly undesirable development and it makes the system very weak.* Usually a few show-pens are available behind the judging table. Promising cats are kept there, less promising is sent back. At the end of the class the 4 best cats (again optional) are presented simultaneously (either by stewards or by the owner). The *simultaneous presentation of the best cats by stewards or owners is an advantage of the continental system*. It enables the judge to compare all parts of the best cats in one glance. Less information has to be retrieved from long-term memory. Then the judge gives them a rank order and decides whether number 1 deserves a certificate.

c. The American or <u>ring system</u> is comparable with the British system in being fast. It is *less vulnerable to the predecessor effect* than the traditional system because all the cats in the group are present in the ring. *Stewards and owners are not involved*, only a clerk. The bias a clerk might cause is comparable with the influence a steward might exert on the judge. The selection of the best cats is similar to the British system, but here the cat is taken out of the pen by the judge him/herself. *A disadvantage of both the ring and British system is the fact that the opinion of a judge is based either on how the cat is sitting or lying in the pen or how the cat presents itself when handled by the judge.* As only one cat at the time can be handled the judge has to rely on his or her long-term recall for a great deal when deciding for the best cat. There is no possibility to see them all in one glance, as the stewards needed for such a simultaneous comparison are not available. As the cats remain in the show-pen when compared simultaneously, *there is a greater chance for judges to become a victim of*

head judging, as body, coat colour, and coat texture are more hidden than the head. No written reports are given. At the end of a group some general remarks on the quality of the cats are verbally expressed and the judge usually comes forward with a more extensive motivation of the best and 2nd best cat.

It is up to the reader to decide *which system is the best*. The outcome will depend upon the *goals set beforehand*. If you want a fast system a choice has to be made between the GCCF and the ring system. Do you want to rely on written judgement reports instead of on verbal explanations then the GCCF system is better than the ring system. In the ring system only verbal reports are given which cannot be verified unless they are recorded. Do you want to give extensive feedback to all the exhibitors, regardless of the qualification of the cat? Then you should choose the continental system. Do you want a more reliable and valid system? Then the continental system deserves disapproval since the introduction of the condemnable habit that owners were permitted to act as stewards. Anonymity of the cats is no longer ensured, which constitutes a serious source for all kinds of biases. Do you want many prizes? Then you have to make a choice between the GCCF and the continental system. Titles or champion certificates are awarded at different levels for colours within races. The continental has almost unlimited possibilities: a cat can climb up the ladder from CAC, over CACIB, and CAGCI, up to CACE. Because of the growing numbers of coat colours classes have become very small nowadays. There is hardly any serious competition within classes. Often a cat is alone in its class. A cat without obvious defects or faults will almost automatically get its title. The more certificates are awarded the happier the exhibitors are, but it should be noted that *lack of competition leads to devaluation of titles*. In some breeds, such as the Main Coon and the Norwegian, coat colour is irrelevant in the standard. Only type, eye form, and coat texture count. Therefore it is illogical to divide both races in colour classes, which now happens all over Europe. This silly rule should be abolished. In my opinion *the Norwegian and Maine Coon standards should be changed in such a way that they are judged according to coat colour as well (just as the Turkish Angora and most other semi-longhairs).* An alternative would be to leave the standards as they are, but to judge the Maines and Norwegians in colour groups (plain colours, silvers, tabbies, bi-colours, etc.). Then there would be more competition, and as a consequence the devaluation of titles would be arrested. At present it could be argued that titles are more or less taken for granted. BIS is what counts in the continental system. In that sense BIS-judging is comparable with judging in rings. In rings there is more competition as well, which makes it more interesting. However there is one huge difference. Only one judge is involved in a ring. For BIS several judges are involved which makes the end result less subjective. How is coped with subjectivity in the ring system if there is only one judge involved? In the same way that colour classes (or open classes in the UK) are judged in the GCCF and the continental system, e.g. *by having the same cat judged on several occasions by different judges*. Hence there are two procedures to reduce subjectivity:

1. Either a panel of judges on only one occasion, such as in BIS, or
2. Each time another judge on several occasions such as in colour classes per breed or when judging groups of different breeds in a ring.

The philosophy behind “each time another judge on several occasions” is related to the *assumption that cats don’t change much over time*. That might be true, although it should be noted that coat length and overall condition are less consistent. However what is more likely to fluctuate is competition from other cats. A cat awarded with a CAC on one show may be beaten on another show. From the foregoing it is concluded that this procedure to reduce subjectivity is far from waterproof.
In other areas of human judgement *where subjects are more likely to fluctuate over time* the “panel of judges on only one occasion” option is preferred. This procedure is used in many aesthetic sports, such as figure skating and gymnastics. Although panel judging in cats (in case of BIS) is less sophisticated than panel judging in aesthetic sports, where advanced statistical methods are applied to compute the final results, all forms of panel judging have in common that they try to cancel out individual deviations from the norm. Furthermore in figure skating and gymnastics (hardly ever on cat shows) the marks are presented to the audience. It is a form of open judging, where it can easily be seen which judge deviates from the norm. If necessary he/she can either be reprimanded or suspended (unthinkable on cat shows!). However, the marks once given cannot be altered.

Towards a more objective system for choosing BIS and Best Cat of the Year

From the foregoing it can be argued that the quality of judging cats for BIS can be improved by using the methods and knowledge derived from other areas of human judgement. Promising candidates are judging systems applied in figure skating (for instance the majority system) or gymnastics. When transplanted on judging cats it would imply that *each judge in the BIS-panel should give a rank order to all the cats proposed for BIS.*
If all judges would give the same rank order to all cats on the podium an optimal *reliability* (concordance) would be obtained (100 % agreement between the judges). This outcome usually means an optimal validity (picking out the best cat) as well. However, it should be kept in mind that *a high reliability does not automatically imply a high validity*. If all the judges become victims of the same bias (for instance choosing the cat of the president of the cat club) the reliability remains high (high agreement between the judges), but the validity will drop, because the president’s cat was not the best cat.
Fortunately this is an extreme example. Generally human beings have a tendency to disagree (not per se intentionally), or to see things 'differently'. Therefore a panel of judges is required (uneven number). If the agreement between judges would always be 100 %, only one, instead of a panel of judges, would be sufficient.
Another advantage of the majority system is that all the cats proposed for BIS will get a rank order. *If the results of all the BIS events are taken together over one year then a more objective method is available to appoint the best cat of the year! The cat with the lowest total is the winner*. However the number of shows should be equal for each cat. In case of ties the rank order on the most prestigious show is decisive.

Concluding remarks

First the most common biases found in cat judging were discussed. Then a comparison was made between the three most prominent judging systems. An alternative method for choosing best in show and best cat of the year was presented as well. However despite all these efforts, meant to harmonise individual bias in judging, we have to accept that the judging of cats resembles other assessment procedures in not having a more objective measuring device for assessing performance than the human eye and brain. We cannot change that. Judging remains subjective. What then is the surplus value of this article? The contribution of this article might be that mere pointing out to the existence of involuntary biases and to the psychological mechanisms underlying them already has a beneficial effect. Without this knowledge involuntary biases in judging cats were unnoticed. It could be argued that they acted as an enemy in disguise. Now the enemy has been unmasked. It goes without saying that it is far easier to beat a visible than an invisible enemy.

Further Reading

This list contains titles for those who wish to read in detail about particular objects mentioned in this article.

Governing Council of the Cat Fancy (1999). The official Standard of Points. Bridgewater: GCCF.

Gregory, R.J. (2000). Psychological testing (3rd Ed.). Boston: Allyn & Bacon.

Kolb, B. & Whishaw, I.O. (1996). Fundamentals of human neuropsychology (4th Ed.). New York: Freeman

Moormann, P.P. (1994). Involuntary bias when judging figure skating performance (Chapter 9). In P.P. Moormann (Ed.), Figure skating performance (p.p. 171-198). Leiden University (ISBN 90-9006862-7).

Pond, G. (1972). The complete cat encyclopaedia. London: Heinemann.

Tabor, R. (1991). Cats: The rise of the cat. London: BBC Books.