Evaluating excellence in the arts

Victor Ginsburgh
ECARES, Brussels and CORE, Louvain-la-Neuve

and

Sheila Weyers
Université catholique de Louvain, Louvain-la-Neuve

January 2013

“I remember my elder colleague Ernst Kris [keeper at the Kunsthistorisches Museum in Vienna, and psychoanalyst] returning from a trip to Italy and my asking him eagerly what new insights about psychology of art he had brought back. ‘I have made a discovery,’ he said gravely. ‘It is the great masters who are the great masters.’” (Ernst Gombrich, 1979, p. 165)

Though many philosophers, historians, psychologists, or sociologists deal with the question of genius in the arts, their definitions hardly help identifying such individuals. For Kant (1790, § 49), genius “is a talent for producing that for which no definite rule can be given [and] originality must be its first property… [its products] must be models, i.e. exemplary.”

A more “operational” definition had been proposed a few years earlier by Hume (1757, p. 9) in his celebrated essay “Of the standard of taste”: “A real genius, the longer his works endure, and the more wide they are spread, the more sincere is the admiration which he meets with.” This definition contains the idea of the “test of time.”¹ To make his point clear, Hume adds that:

“envy and jealousy have too much place in a narrow circle; and even familiar acquaintance may diminish the applause due to [the artist's] performances: but when these obstructions are removed, the beauties immediately display their energy; and while the work endures they maintain their authority over the minds of men.” (1757, p. 9)

Literature Nobel Prize winner John Maxwell Coetzee (2002, p. 18) argues that

“the criterion of testing and survival is not just a minimal, pragmatic standard. It is a criterion that expresses a certain confidence in the tradition of testing, and a confidence that professionals will not devote labour and attention, generation after generation, to sustaining [artworks] whose life-functions have terminated.”

¹ See Savile’s (1982) impressive work for a thorough discussion of the concept.
In other words, artists (painters, sculptors, architects, writers, composers, music interpreters, actors, etc.) and masterworks (including collective creations, such as movies, see for instance Simonton, 2004) that are judged important also have economic value, and agents will be ready to incur costs to have these works surviving in good condition, and the names of their creators will be remembered.

However, if according to Kant, exemplarity and originality characterize genius, he also claims that

“one cannot describe scientifically how genius brings about its products. A genius does not know how he has come by his ideas and cannot formulate precepts which will enable others to produce similar works... Artistic creation is a mystery that no amount of explanatory investigation will dispel.” (White, 1992, p. 90.)

Likewise, in the introduction of his celebrated Story of Art, Gombrich (1972, p. 17) writes that it is usually impossible to express in words what makes us recognize a masterwork, which, nevertheless, does not mean that there is no hierarchy. It thus seems less demanding to avoid the difficulty of pointing to geniuses only and consider what art historians, art philosophers and sociologists came to call “canons,” for which we adopt the definition given by Camille (1996, p. 200): A canon is of a set of objects or artists that “stripped of contingency, transcend space and time and stand autonomous.”

Though canons evaluate by isolating who is “in” from who is “out,” they do usually not formally rank, first, second, ... But mankind has for very long also been fascinated by competitions and rankings, within or without a canon. According to Huizinga (1951), both competitions and rankings are close to games that seem to be even older than mankind. He shows that all the characteristics of gaming are already present among animals: some birds decorate their nests, some produce melodies to be appreciated by others of their kind, and even crows ‘organize’ flying competitions. Therefore, competitions as entertainment do not seem to be caused by culture, but predate it. They are present in the most futile games, but can also lead to mortal fights, that are unfolding according to well-established rules. In ancient Greece, men pit themselves against others in order to win, whatever the cost. Olympiads, the first of which was organized in 776 B. C. is one such example.

Peisistratos, the tyrant who ruled Athens in the sixth century B. C., was at the origin of theatre contests organized every year between tragic Greek poets. The ten randomly
drawn judges often “nominated” Aeschylus, Euripides and Sophocles, who are still part of today’s theatre canon.

In chapter 36 of his *Natural History*, Pliny the Elder reports on a contest between two Greek painters Zeuxis and Parrhasius active during the fourth century B. C. This is how Pliny describes the event, at the times during which imitation of nature was the objective:

> “Zeuxis, who [had] represented some grapes, painted so naturally that the birds flew towards the spot where the picture was exhibited. Parrhasius, on the other hand, exhibited a curtain, drawn with such singular truthfulness, that Zeuxis, elated with the judgment which had been passed upon his work by the birds, haughtily demanded that the curtain should be drawn aside to let the picture be seen. Upon finding his mistake, with a great degree of ingenuous candor he admitted that he had been surpassed, for that whereas he himself had only deceived the birds, Parrhasius had deceived him, an artist.”

The notions of genius, canon and ranking are hard to disentangle. A canon is often a set that contains geniuses. Artists who belong to a canon can be ranked, and the first can possibly be considered geniuses.

The chapter is organized as follows. In Section 1, we discuss canons, how they form, whether they are “ideologically” pure, and whether they are stable or not. We also discuss the difference between open and closed canons, and the idea that canons may consist of concentric circles in which canonicity is negatively correlated with distance to the center. We finally describe how canonic choices are made. Section 2 turns to rankings that result from compilations of encyclopaedias and art history books or from competitions and voting procedures. Section 3 is devoted to concluding comments.

**CANONS**

The first (informal) canon that comes to mind is the one established in the third century B. C., a list of the “seven wonders (actually sites) of the ancient world.” These consisted of constructions (the Great Pyramid of Giza, The Hanging Gardens of Babylon, the Mausoleum of Halicarnassus, the Colossus of Rhodes, the Temple of Artemis at Ephesus, the Lighthouse of Alexandria and the Statue of Zeus at Olympia) visited mainly by Greek travelers during the last centuries B. C. Though each of them was considered unsurpassed in the “country” in which it had been erected, they were not ranked within the canon. None of them was considered of “higher” importance than any other.
There are also examples that are closer to our times, such as the set of individual artists who were selected by Lampsonius in his *Pictorum Aliquot Celebrium Germaniae Inferioris* (1572), which consists of a series of 23 portraits of artists active in the Low Countries, in no specific order. Again, there is a selection, but no ranking. Vasari’s *Vite* (1568) of Italian artists from the Renaissance follows more or less the chronological order. So does van Mander (1604) in the very first history of painters from the Low Countries. In his *Abrégé de la Vie des Peintres* de Piles (1649) describes the lives of painters region by region, and chronologically within each region. Some “lives” are more eulogistic than others, but using hermeneutics to rank artists has, to our knowledge, never been considered.

In his *Cours de Peinture par Principes* (1708), the same de Piles includes a table that he calls *Balance des Peintres*, in which he grades 56 deceased painters on each of four properties (composition, drawing, color and expression) that he finds important. Rembrandt, for example, is very low on drawing and obtains 15, 6, 17 and 12 on the properties just mentioned, while Michelangelo is very high on drawing, with scores of 8, 17, 4 and 8 respectively. De Piles stops short of computing any sort of average of the grades and does not rank his painters, though they probably would all belong to his canon. The details of this interesting as well as surprising exercise are displayed in Table 1.

[Table 1 approximately here]

Who has the right or the power to judge? Hume (1757, pp. 17-18) considers, and most art philosophers and critics concur with him, that quality assessments should be left to experts who are familiar with the experience of works of art:

“though the principles of taste be universal... yet few are qualified to give judgment on any work of art, or establish their own sentiment as the standard of beauty... Some men in general, however difficult to be particularly pitched upon, will be acknowledged by universal sentiment to have a preference above others,”

but the question of how to find such people is, as he writes “embarrassing.” Nowadays, art critics and historians, art dealers, Unesco’s World Heritage list, art historical manuals and surveys, public art collections and prominent museums obviously participate in the formation of canons. There is, for example, little doubt that the New York Museum of Modern Art will have its word on the canon of twentieth century Western art.
One or several canons

Assessing whether an artist, a sportsman, a politician or a scientist is a genius, or whether, somewhat more weakly, he belongs to a “canon,” that is to an exemplary set of people or objects (artworks or texts), needs several decisions, and poses several questions, the first being “to which set.”

This idea is very well illustrated by William C. Rhoden, a sports columnist for The New York Times, who wonders whether Michael Phelps could become the “greatest Olympian of all times.” Here are some excerpts of the opinion he expressed after the 2012 Olympic games:

“Michael Phelps will leave London as the most decorated athlete in Olympic history. The raging debate is whether this does make Phelps the greatest swimmer of all time, the greatest Olympian of all time or the most successful Olympic athlete of all time. One thing is clear, Phelps is definitely successful and clearly the greatest and most consistent swimmer of his generation. But the greatest Olympian and still champion is the great Jesse Owens, the American sprinter and long jumper who became an international hero at the 1936 Berlin Olympics. Yes, Phelps has won more medals than Owens, lots more, and Olympic sprinters have run faster, lots faster but no one has faced down history and performed with as much grace and international scrutiny as Owens did in Berlin. And more than seven decades later, Owens legacy continues to resonate across generations. No, not because he ran fast and jumped far, but because he stared down Adolf Hitler and he did it by allowing his actions to speak loudly to the world. Phelps made Olympic history with his mountain of medals [but] enduring greatness means having an impact not simply on one sport but having an impact on humanity and in that respect Owens defeats Phelps as the greatest Olympian by the length of history.”

Though this quotation is devoted to sports, the question applies to other fields as well. There is obviously no unique canon in art or science, and one has to decide carefully on a case-by-case basis to which canon an individual or an object should belong. Art in particular is culture-dependent and it would be heroic to introduce the Japanese painter and printmaker Hokusai into the same canon as Rembrandt, who used exactly the same media as Hokusai. And so would it be to construct a unique canon, even for such a short period as the Renaissance, in such a small part of the world as Europe, which would contain names of both Italian painters and their Flemish contemporaries.

Though they were very often in contact and visited each other, their styles and manners are very different. Great scientists are much more likely to be accepted as universal geniuses, but one would still hesitate to define a single canon containing mathematicians and biologists. We leave it at that with Gombrich’s (1979, p. 162) opinion:

“It would be a real impoverishment if we demanded of all masterpieces that they should coincide with our own value system. It would be an impoverishment precisely because art is not life, and can help us extend our sympathies and our understanding of basic human reactions, which transcend the limits of any one culture or value system.”

*Canons and ideology*

Meanwhile, canons are tainted by other considerations than the assumed beauty and trans-historically stable value of its elements. Locher (2012, p. 37), for instance, mentions that

“canons are more or less collectively developed and agreed reference systems, representing sets of values deemed to be important for society as a whole, or for groups within it; [as a consequence] canon formation has to be considered as a social and political enterprise.”

Canons have indeed become the object of fierce discussion in the late twentieth century, as exemplified by the following (Waugh, 1992, pp. 59-61):

“The development of feminist, Marxist, postmodernist and post-colonial approaches to art, the sense of increasing commodification of the aesthetic in Western societies, of new artistic forms arising out of technologies such as video, computers and television have all contributed to a new ‘explosion’ of the traditional identities of high art. These changes have given rise to a variety of arguments which represent the canon as an ideological formation bound up with relations of power within those institutions which are seen to regulate cultural values and notions of taste… The canon is seen to function as an instrument of exclusion through the construction of a value system which legitimates as good those artefacts which mediate or represent the identities of those with cultural power.”

This echoes the French sociologist Bourdieu (1983, 1996) who argues that evaluation, and thus value, is arbitrary, because it is based on motivations imposed by the social and political structures of the cultural hierarchy. It is objective but only as a social fact: the artistic field is contained within the field of power, which is itself situated within the field of class relations (Bourdieu, 1983, p. 319). Accordingly, there are no criteria
that allow determining the intrinsic quality of a work, but only professional judges who “possess the socially accepted authority to ascribe specific properties to a work ... and how it should be ranked” (Van Rees, 1983, p. 398. See also Rajagopalan, 1997, and Van Peer, 1996).

Writing about literary canons, von Hallberg (1984, p. 14) is not optimistic either. The canon, he writes, is

“the traditional dream of ambitious critics. A canon is commonly seen as what other people, once powerful, have made and what should now be opened up, demystified, or eliminated altogether;”

which is close to Westphal (1993, p. 440) who considers that “establishing a canon is a ‘violent’ act because it institutes an exclusionary hierarchy with social power but without ultimate justification.”

*Open and closed canons*

A question that is related to the previous one is whether one looks at “closed” or “open” canons. We define a closed canon as being a set of objects or artists that can no longer grow, because nothing is created that can be added. For instance, the canon of Italian Renaissance painters is essentially closed, unless one discovers a painter who is not yet known, which is fairly improbable.

Alternatively, a canon is open, if the production of works, and of those who produce them keeps increasing. The canon of American movies, and of their directors, actors, etc. is still open to new works. The canon can thus keep growing if no limit is put on the number of elements that the set contains. Alternatively, if the number of elements is fixed, instability may follow. This is so in the British movie industry magazine *Sight & Sound* canon that the film critic and screenwriter Roger Egbert considers “by far the most respected of the countless polls of great movies—the only one most serious movie people take seriously.” The canon results from a poll conducted every ten years (since 1952) among a wide and international group of film professionals.

Table 2 shows how the canon behaves over the 1952-2012 sixty-year period. Only one film, Renoir’s *La Règle du Jeu*, remains in the canon over the whole period; two celebrated movies, *Potemkin* which entered in 1952 and exited in 2012 and *Citizen Kane* (1962-2012) are present during 50 years, *Vertigo* is there since 1982, while 22
movies spent only 10 or 20 years. This shows that in open cannons, large swings are likely to occur.

It is also interesting to note that experts choose movies that already seem to have passed the test of time: 19 out of the full list of 32 films that entered, stayed or exited were produced before 1952, the year in which the first list was established. Thirty-one films out of 32 were produced before 1950 and only one (Godfather I and II) was produced after 1950; it entered in 2002 only, and left in 2012. The youngest movie in the 2012 canon, Kubrick’s 2001, A Space Odyssey, was produced in 1968.

[Table 2 approximately here]

Formation of a canon: How do artists (or works) attain canonical status?

In her paper on canon formation, Silvers (1991, p. 212) suggests that

“understanding how evaluative critical judgment evolves might be supposed to require detailed empirical study more properly pursued in disciplines other than philosophy because the nature of the process which forms canons is sociological, political or economic, rather than autonomously aesthetic.”

She describes several paths that may lead to canonical status. According to her, a work (or an artist) may qualify in one of the following three ways:

(a) Failing, despite systematic scrutiny, to reveal defects or disagreeableness sufficient to be disqualified, or

(b) Revealing previously unnoticed meritorious or agreeable properties sufficient to qualify, or

(c) Acquiring valuable properties sufficient to qualify it.

She points out that (a) and (b) account for this process in terms of events that occurred before or during the life of the artist and possible changes in the opinions of art scholars (“traditionalism”), while (c) accounts for events that took place after a work was produced and that change its properties (“revisionism”).

According to traditionalist art theorists, all properties are present when the work is conceived and realized, though their importance may have been overlooked. Leonardo
da Vinci’s oeuvre is an example that satisfies criterion (a). The *Grove Dictionary of Art* (1996, vol. 19, p. 196) suggests that “there has never been a period in which Leonardo’s greatness has not been acknowledged.” Criterion (b) can be invoked when attributions are revised. Duccio, for example, appeared in the canon after Berenson reattributed to him the *Madonna di Ruccellai*, previously thought by Cimabue. In such cases, changes in canonical status are due to epistemic reasons.

Criterion (c) is a form of *revisionism* according to which some properties or attributes may be discovered in existing works by newly created works, or because earlier works “acquire salience in relation to the proper understanding of [new works], which they naturally didn’t have before” (Levinson, 1996, p. 268). Silvers (1991) illustrates what happened with Rubens’ figures after Renoir and Picasso. Rubens himself considered his figures to be coarse, and this was accepted for centuries after he painted them. Today, writes Silvers (1991, p. 217):

> “when art’s history embraces treatments of the human figure such as those painted by Renoir and Picasso, Rubens’ treatments are transfigured to become fluently refined and elegantly vital… the composition of Rubens’ paintings remain[s] the same, but the works’ aesthetic attributes change, develop, transmogrify or evolve.”

Junod (1995) similarly suggests that Vermeer was rediscovered in the mid-19th century due to the closeness of his work to pre-impressionist sensitivity. Whether such situations change the properties of the work itself is hotly debated among art philosophers. Levinson (1990, p. 194) for example argues against revisionism, and brings the revisionist argument to an extreme with another example: After Cubist painting came into existence, “the non-Cubist mode of depiction of, say, Holbein’s *Ambassadors* suddenly appears as an artistically relevant attribute of this painting,” though the painting was produced in 1533.

Both traditionalism and revisionism are present in the narratives of art historians who, by continuously provoking our attention, are among the most important contributors to the formation of canons. That traditionalism is present is obvious, since historians evoke what occurred *before and during* the creation of the artworks that they describe. Revisionism comes into narratives because art historians also take into account what happened *after* the creation of the works. Discussing Goya’s influence on Bacon in an article devoted to Bacon qualifies Bacon. But Bozal’s (1997) observation in his monograph on Goya that the “horribly open devouring mouth in *Saturn* is a prelude to the howling mouths of Bacon,” may be an addition, even if only second-order, to the fame of Goya. However, one can argue whether this changes Goya’s *Saturn* itself, or
whether it merely changes our vision of the work. Therefore, it is often difficult to decide whether a work qualifies by criterion (b) or (c).

Note that the above analysis of canon formation implies that artworks are endowed with properties, a view that is accepted by art theorists, though they do hardly agree on the set of properties.

Ginsburgh and Weyers (2010) follow Silvers’ idea and analyze the formation over time of the late 20th century canons of two schools that dominated all other European schools in their time: Italian Renaissance and Flemish Realism. Since most artists were discussed, some 400 years ago, by Vasari in the second edition of his Vite, and by van Mander in his Schilder-boeck, narratives by art historians can be followed over a long period of time. To explore the dynamic process of canon formation, the authors collected data on the presence and the greatness of a large number of artists in narrative works written by important art scholars at time intervals of roughly 75 years, so that the 400 years elapsed between 1600 and 2000 are spanned as best as possible. Length of entries are used as a measure of greatness—see more on this later—and the canon in each period is defined as consisting of the 50 names that have the largest entries. They conclude that at least half of the fifty artists who appear in the late 20th century canons had already been introduced 400 years ago, and satisfy Silvers’ criterion (a).

Wrong attributions or new technical discoveries prevented some names to be canonical any earlier. Some artists entered the canon at later times, given that art historians learned to appreciate or to understand them better. This corresponds to Silvers’ criterion (b). The group of artists who entered or were moved to the first circle because their works acquired new properties in the light of later works, as suggested by Silvers and her criterion (c), is small.

Though criteria such as invention, originality, newness and progress, and their relative weights in evaluating artists change over time, it is surprising that half of the names appearing in the two canons were there almost from the beginning. This appears to be in contradiction with the suggestions made by Junod (1976), Genette (1994) and many others that canons are continuously changing and that no artist can survive forever. One reason for this apparent discrepancy may be that here, one studies two “closed” canons, that is, canons that make no room for artists who were born after 1600, and for rather small (even if artistically important) regions of Europe. A contemporary canon that would be devoted to “European great painters of all times” would probably
include Manet, Duchamp, and Picasso, and maybe exclude Botticelli. But Leonardo, Michelangelo, Van Eyck and Rubens would probably be there to stay forever.

The closed character of the canons that we examine, as well as calling on well-known historians may explain why there are less discoveries or rediscoveries and shifts than those described by Haskell (1976). His celebrated work is mostly based on the behavior of art collectors and much less on the opinion of art historians, and changes of taste to which well-trained art historians such as those on whom Ginsburgh and Weyers based their research should be less prone. Botticelli never left the canon. He became indeed peripheral during 200 years after Vasari’s description, but regained centrality with Burckhardt (1855) in his *Cicerone*.

In short, there are several cases of wrong attributions or of new technical discoveries that prevented some names to become canonical earlier, though there are also artists whom art historians learned to appreciate later. Finally, the number of names that entered the canon because their works acquired new properties in the light of works by artists that followed them is not very large.

The observation that canons are reasonably stable, is also present in several other studies. Milo (1986) shows that French painters who were active during the 16th and 17th centuries are still praised during the 20th century. Landes (2003) looks at some 850 American artists active in the late 19th and early 20th century, whose paintings were shown in three important exhibitions in Paris in 1900, and in New York in 1913 and 1939. He tests whether those who were recognized by awards, or elected as members of a distinguished academy have more staying power than others. Rosengren (1985) applies a similar idea to study how the fame of Swedish writers born between 1825 and 1849 is perceived in two points of time (the 1880s, and the 1960s). Verboord (2003) also describes such procedures to assess the value of some 500 writers, using indicators such as awards, the number of academic studies devoted to each author, literary encyclopedias, the literary status of publishers. In all cases, there is more stability than expected.
Stability of the canon

The various configurations are cast in a statistical framework by Simonton (1998), who offers the following useful classification:

(a) *Transhistorical stability.* Successive generations may disagree over time, but not in a systematic way, suggesting that they apply largely the same set of criteria in evaluating works.

(b) *Exponential decay.* Judgments by a generation take into account judgments of the immediately preceding generation, suggesting that the sequence is governed by a first-order autoregressive process, implying decreasing correlations of evaluations over time.\(^3\)

(c) *Gradual attrition* or steady decline, in which case correlations between contemporary and subsequent judgments would decline in a linear way.

(d) *Cyclical fashion,* with periodic or quasi-periodic fluctuations in assessments.

(e) *Complete transhistorical instability,* if judgments lack any consistency over time. This leads to zero correlations of assessments over time.

Simonton (1998) studied 496 operas composed between 1607 and 1938 by 55 composers. He finds that (i) the success of an opera is a positive monotonic function of its success when it was created for the first time, (ii) the consensus is not stable over time and (iii) the instability is due to cycles. He elegantly concludes that “for the creative genius, perhaps, immortal fame may simply skip along the surface of fickle fashion, deftly leaping from peak to peak and gently gliding over the deepest valley.”

Ceulemans’s (2010, p. 237) who deals with the reputation of 366 Baroque music composers over time shows that musicologists rely on their predecessors’ work, and concludes that cultural tradition matters. Using some sophisticated econometrics, he suggests that “composers do not appear and disappear suddenly, they rather move up or down smoothly, but relatively fast.”

Ginsburgh and Weyers (2006a) look at all the painters to whom Vasari (1568) devoted a *vita,* and like Simonton, follow their reputation over time, using the lengths of entries

\(^3\) Note that this needs the coefficient of the autoregressive process to be smaller than 1.
in seven art histories, beginning with Vasari and ending with Turner’s (1996) *Grove Dictionary of Art* (1996). They observe that though some artists appear, disappear or reappear, there is a surprisingly large degree of consensus over time: among the first 50 to whom scholars devote space, one half is recognized at all times, an observation sustained by several statistical tests.

*Circles of a canon*

Westphal (1993, p. 436) develops the idea that canonicity is

“not an either/or matter, but a matter of degree. At any given time, the canon is best represented by a series of concentric circles. At the center are [works or artists] with the highest degree of canonicity, while at the periphery are those whose classical status is most tenuous… This means of course that historical changes in the canon are not simply matters of inclusion and exclusion, but also matters of location between the center and the periphery.”

Elements (works or artists) can thus move between circles over time, whether the canon is open and newly created works can enter (movies), or closed (the Italian Renaissance). This way of looking at canons is a smooth introduction to what will be discussed in the next section concerned with rankings of artists, in which each element of the set can be considered contained in a circle.

This view is also emphasized by Cutting (2006, p. 13):

“there is no sharp distinction between canon and the broader corpus that could be conceivable for canonization; the former contains a graded hierarchy with some works primary, other secondary, still others tertiary, and so forth, until one reaches the base corpus.”

And this is precisely what he tries to do in Cutting (2006, Chapters 7 and 8).

The number of artists who belong to a canon is of course also arbitrary. Gombrich (1979, p. 157) argues that Beethoven’s slow movement of his String Quartet op. 132 will never “belong to the top ten, [b]ut it does belong to the canon, and rightly so,” and the canon could very well contain 11 composers.

*Canons in practice*

Canons are usually informal or implicit. Art historians cite masterpieces and names of

---

4 Westphal deals with the canon for “texts”, but what he writes extends easily to any other artworks or artists.
artists, but they rarely set up “lists.” Gombrich’s Story of Art (1972) that covers artistic creations from prehistoric cave paintings to 20th century art contains a chronological table with names of artists (Dürer, or Morandi) as well as masterpieces (the Temple of Zeus at Olympia or the Alhambra in Granada) but there is no pretension by Gombrich to call this “canon.” Of course one could constitute a canon by collecting the names of artists and/or masterpieces cited by Gombrich.

Canons may also result from choices made by experts, such as the collection of Great Books of the Western World, of which the first edition of 54 volumes (from Homer to Freud), resulting from the collaboration between the University of Chicago and the Encyclopaedia Britannica, was published in 1952. The collection has been reedited in 1990. Another example consists of the so-called World Library compiled in 2002 by the Norwegian Book Club, and based on choices made by writers each of whom had to select his list of ten books. It contains 100 books from all countries, cultures and periods, but no ranking, with the exception of Don Quixote considered the “best literary work ever written,” in some sense the “genial” work. Other literary canons result from polls conducted among readers, such as the 100 Books of the 20th century compiled by the French leading newspaper Le Monde. In other cases, such lists are drawn by a single personality: Marcel Reich-Ranicki, considered as the best expert of German literature edited an anthology of more than 50 volumes of German novels, short stories, dramatic plays, and poetry. And though many names are by now forgotten, the list of Nobel prizes in literature could also be considered a canon.

Movies are also a field that is prone to generate “100 best movies lists,” some of which are the work of a team headed by well-known expert (Leonard Maltin), others are set up by several movie experts or by moviegoers.

Most recent studies in the arts use historiometrics (see Simonton, 2007) that start with the collection of representative samples, or surveys of the entire population in a given field, such as all painters from the Italian Renaissance, or all Baroque music composers. They proceed by retrieving variables that describe or summarize the career and creations of each individual such as the number of images in art history books (Cutting, 2006, Galenson, 2001, 2002, Ginsburgh and Weyers, 2006c), the number of exhibitions, the number of prestigious collectors or museums that own their paintings (Cutting, 2006), the number of times an opera was performed or recorded (Simonton, 1998), or the length of entries or number of citations devoted to painters or composers.

---

5 The series has its own website at http://www.berkanon.de
6 See http://www.filmsite.org/greatlists.html

**RANKINGS**

In some sense, even if individual artists who appear in the various canons are not necessarily ranked, they are sometimes so because scientists set up canons by counting lines, images, number of citations, etc. which *ipso facto* lead to rankings, or to concentric circles. Art historians and art philosophers usually discredit such methods. Even de Piles (incomplete) ranking evoked earlier was not taken seriously. He himself looked at it as a game, though his contemporaries considered it an “ingenious way to characterize genius” (Thuillier, 1989, p. xxvii). Later on, this view changed. Von Schlosser (1924) an important member of the Vienna School of Art History, hated it. Junod (1996, p. 505), a Swiss art historian, thinks that “[it] pretty much looks like the prize-lists set up by some of our contemporary art critics.” In his book on de Piles’ theory of art, British art historian Puttfarken (1985, p. 42) thinks of him as having been “at his worst when he tried to be most systematic.” And the greatest of art historians, Gombrich (1966, p. 76) judges it a “notorious aberration.” He gets a little gentler with de Piles later on but still his advice is radical (Gombrich, 1979, p. 153): “The way a great artist collects his points is not susceptible to quantification. Roger de Pile’s brave attempt to do precisely this should have established this impossibility for good and all.”

French art historian Thuillier is more positive. Though in his preface (1989, p. xxvii) to a recent edition of de Piles' *Cours de peinture*, he compares the *balance* with “contemporary art dictionaries which set to three pages, half a page, fifteen or five lines the length devoted to each artist, and make it an unpleasant exercise” he adds that this exercise of “unpleasant” quantification may perhaps nevertheless “be necessary.”

Others, such as Teyssèdre (1965, p. 187), the French art historian and expert of de Piles, supports quantification: “even if ratings are difficult to assess, one could just look at whether artists are cited or not.” Robert Rosenblum (2002) reminds us that the British art critic and curator, David Sylvester

“could silence the cheerful gossip at any dinner table by posing with urgent solemnity a question like ‘But who do you think is greater, Giotto or Matisse?’ [Even] as he lay dying, he insisted on having his guests play his favorite game of quantification, listing who are the greatest of the great.”
In what follows, we discuss two issues:

(a) How can one aggregate ratings or rankings made by a jury, where the jury is either a group of people of which each member has to rank participants in a competition, or a scientist who uses books or other repertories to count citations, images, or number of lines.

(b) If one agrees with the idea that artworks (and artists) carry properties each of which can be rated, how can one aggregate the ratings and end up with a summary rank or rating for each work or artist.

*Aggregating ratings or rankings produced by a jury*

We assume that $m$ judges (books, lists, real judges) mark or grade $n$ objects (here artworks or artists). The marks obtained are then aggregated (usually added) and the outcome results in a unique ordering (though there may be ties). There are two problems with ranking on the basis of marks. Some judges are generous and give high marks (or write thick books, or use many images) others are less so. There also exist judges who scatter their marks more than others, and by doing so are able to “express stronger preferences by numerical differences” (Ashnefelter and Quandt, 1999). This shows that one has to normalize marks to make them comparable, and the method used is to transform the marks into ranks that can be aggregated (or added), to produce the final ranking.

To show, however, that this method may not be satisfying either, consider the following example of 3 painters A, B and C and 3 judges. Judge 1 ranks A first, B second C third; judge 2 rank B first, C second, A third; and judge 3 rank C first, A second, and B third. Consider now pairs of candidate painters: A is preferred to B by 2 judges (1 and 3); B is preferred to C by two judges (1 and 3). So obviously, if A is preferred to B, and B is preferred to C one should find reasonable that A is preferred to C. But as can be checked, C is also preferred to A by two judges (2 and 3). Therefore it is impossible to rank them, though each judge provided a ranking. This is an example of the Condorcet Paradox, which shows that there is no “winner,” since every painter is preferred by two of the three judges. Perhaps all three should be considered geniuses…

I is also a special case of Arrow’s Impossibility Theorem that can be (loosely) stated as follows:
When there are at least three choices (here painters, or artworks), there is no aggregate ranking that can simultaneously satisfy the following four reasonable axioms:

**Axiom 1. Unrestricted domain.** All individual preferences are allowed, that is each judge can choose his ordering of A, B and C.

**Axiom 2. Pareto efficiency.** If every judge ranks A before B, then the aggregate order must rank A before B.

**Axiom 3. Independence of irrelevant alternatives.** If A is ranked before B, then introducing a new choice C (or discarding a choice C from the list of choices) must not make B ranked before A: C is irrelevant in the choice between A and B.

**Axiom 4. Non-dictatorship.** No judge can impose his own ranking.

If one accepts that a ranking should satisfy these axioms, one has to accept also that it is impossible to devise a method that can produce an aggregate ranking. Therefore, as Dante once said “Abandon hope all ye who enter here.” But there is another problem, since judges have the difficult task to state their preference ordering over the full set of items. Unless, as mentioned earlier, one counts the number of lines in an encyclopaedia, or the number of citations or of images or of performances of operatic works, it is very demanding to ask one or several judges to rate or rank a large number of items, whether these are artworks, painters or skaters who participate in a competition.

Borda (1781) suggested using approval voting (Weber, 1978, 1995 and Balinsky and Laraki, 2010) in which each judge can cast a vote for as many items as she wishes, without ranking. The votes are then added item-by-item, and lead to an aggregate ranking. Ginsburgh and Zang (2013) suggest a variant in which, if a judge votes for a sub group of size $k$, $0 \leq k \leq n$ of the $n$ items, then each of them gets a fraction $1/k$ of one vote (a judge who chooses $k = 0$, does not vote). These fractions are then added as above, and a ranking is computed. Though both procedures look simpler than ranking, since they only require to choose or not to choose an item as “meritorious,” they do not escape Arrow’s impossibility theorem. However, both methods ask each judge to provide a partial ordering, while the method used in most cases (competitions, for example) requires each judge to provide a complete ordering. Both the Borda and the Ginsburgh-Zang methods lead to a unique ranking (possibly with ties), but the second one also satisfies some reasonable axioms imposed by Shapley (1953) to obtain the so-called Shapley Value, which measures the “power,” “influence,” or “weight” of each judge.
item. (For a proof of this result, see Ginsburgh and Zang, 2003). And this is exactly what we want a ranking to achieve.

Balinski and Laraki (2010, pp. 129-159) devote one chapter of their book to judging in several types of situations and competitions (musicians, gymnasts and skaters, divers, wines), and several chapters to grading systems. The usual system is to grade or rank all candidate items or people, though if this number is too large, or not well defined (for instance all movies produced since 1920), each judge is asked to provide his choice of say, ten items or names and those cited by the largest number of judges are ranked.

Rankings are however often flawed, as they are influenced by other considerations than talent, and prone to short-run considerations influenced by fashion. Flores and Ginsburgh (1996), and Ginsburgh and Van Ours (2003), show that in musical contests, the final ranking may depend on the order of appearance of competing candidates. Ginsburgh (2003) and Ginsburgh and Weyers (2011, 2012) point to the defects of grading and ranking of literary prizes and movie awards.7 Between 1929 (the year the Academy of Motion Pictures started distributing Oscars) and 1995, the number of years in which the Oscar for Best Movie is of better quality (here, quality is measured by the number of “100 best movies lists” in which each movie appears out of a total of 15 compiled at a later point in time) than a nominated movie is equal to 27 (out of 68). There are 10 years in which there is a tie, and 31 years in which the “best nominated” movie beats the Oscar. Awards and prizes are thus very often poor predictors of staying power. Only the test of time provides guidance to establishing canons. Note that this is far from being so in the case of the arts only.

Aggregating ratings of properties

Philosophers, starting with the Greeks, agree that artworks contain properties that contribute to their value. Levinson (2003, p. 6) quotes some properties that belong to an open-ended list: Beauty, ugliness, sublimity, grace, elegance, delicacy, harmony, balance, unity, etc., but adds that the “demarcation of the class is subject to dispute.” Beardsley (1958), Vermazen (1975), and Dickie (1988, chapter 9) take the idea a step further by suggesting that these properties can be rated. However, they also point out that they are often incommensurable, and can, therefore, not be aggregated to compute a total value and make it possible to compare individual artworks or artists, a case that we already encountered with de Piles’ Balance de Peintres shown in Table 1.

7 See also English’s (2005) amazing book on awards and prizes.
To simplify the discussion, we restrict our attention to works that share the same properties and show, using a very simple example that it is not always possible to rank works, even if there is a unique judge (which avoids entering into Condorcet and Arrow paradoxes and impossibility results). Assume that we want to rank two works $a$ and $b$ endowed with the same three properties A, B and C. Numbers between 0 and 20 represent the marks given to each property. Hence $(17, 19, 18)$ means that work $a$ gets 17 on property A, 19 on B and 18 on C. If the properties are incommensurable, one cannot compare this work with $b$ with rates $(17, 18, 19)$, since it gets a lower rate on property B and a higher one on C than work $a$.

Dickie (1988, pp. 167-182) suggests to construct tables that allow comparisons with respect to a given work, say $c$, as long as the works to which $c$ is compared have more of one property, and not less of any other, or less of one property and not more of any other. For example, work $c$ endowed with properties $(16, 15, 17)$ is “better” than the four works that are located below $c$, and worse than those located above $c$ in the following table:

\[
\begin{align*}
(17, 19, 18) \\
(17, 18, 18) &- (16, 19, 18) \\
(17, 15, 18) &- (16, 16, 17) \\
\text{work } c &\rightarrow (16, 15, 17) \\
(15, 15, 15) &- (16, 14, 17) \\
(15, 13, 17) &- (16, 15, 16)
\end{align*}
\]

Such a table contains only a partial ordering, not a complete one, since we cannot decide whether a work with rates $(16, 17, 14)$ is better or worse than $c$.

Dickie (1988, p. 180) claims that “there is no better way or even any other way at all to arrive at reasonable specific evaluations,” and recognizes that critics do indeed rely on such (partial) orderings, though their evaluation is intuitive, and made without any formal construction of the tables that he describes.

There are indeed other ways than the one discussed by Dickie. If properties can be ranked according to their order of importance (lexicographic ordering), then an overall value can be computed. The property considered most important is looked at first and works are ordered according to the rates of this first property. In case of ties in the first property, one orders according to the rates of the second most important property, etc.
For example, if property A is always “more important” than property B, that is itself more important than C, then work $a$ rated (18, 16, 16) would be overall more valuable than work $b$ rated (17, $x$, $y$) where the ratings $x$ and $y$ of properties B and C can be any number between 0 and 20.

Another obvious scheme is to give equal weights to all properties, so that the overall value results from the arithmetic mean of individual ratings. But one can imagine more subtle weighting schemes, and, if art critics are able to decompose works into properties and rate each property for each work, then they should also be able to weigh the importance of the properties and obtain an overall value, even if this last step is informal. This is so when students are rated on the basis of several exams, in musical and other artistic competitions, and in sports such as figure skating, diving, or gymnastics, for which there are several criteria to be taken into account, that do not necessarily have the same importance.

This is very close to the idea expressed by economists, in particular, Kevin Lancaster (1966), according to whom a commodity can be thought of as a bundle of characteristics (or properties). Since there are no markets for individual characteristics, these cannot be bought but consumers can chose among varieties (of cars for instance) to construct their preferred choice(s). Therefore producers (and artists) make these choices for them, and consumers pick the commodity (or the work) that provides the characteristics they find closest to their preferred combination.

For consumer goods, there exist prices that give an (economic) indication of their total value. Economists have also thought of retrieving the value of individual characteristics by using an econometric method (hedonic regression) which consists in regressing prices on the “quantities” of each characteristic (for laptop computers, these characteristics would be size of the memory, speed, dimensions, weight, etc). The regression provides estimated parameters that measure the “weight” with which each characteristic contributes to the price.

This method could be used to derive the weights of the properties in artworks (or artists), but one needs to have, as is the case for consumer goods, an indication of the total value of artworks. This is what Ginsburgh and Weyers (2006c, 2008) tried to do for de Piles painters and for Best Movie Oscars.

In the paper on de Piles, they suggest taking two contemporary measures of total value. One is economic and consists of auction prices obtained between 1977 and 1993 for
artworks of painters who appear in Table 1; the other is art-historic and is based on the length of entries for those painters in the Grove Dictionary of Art (1966). They find that the largest weight is captured in both cases by the property that de Piles calls “color.” “Drawing” has no significant impact, while “composition” and “expression” come out with a significant effect on prices but not on art-historic values.\footnote{See also Graddy (2012) who obtains similar results for prices in a much larger sample over time.} There are two points worth insisting on. First, it is quite surprising that the 300 years old de Piles’ properties and ratings still seem to correspond to what buyers (and art historians, to some extent) take into account today. Second, the weights that are represented by the estimated regression parameters are quite different (0.047 for composition, 0 for drawing, 0.112 for color and 0.046 for composition). A simple arithmetic average of rates (that is, equal weights) would thus not be appropriate.

In their paper on movies, Ginsburgh and Weyers (2008) try to verify whether Oscars or nominations for actor or actress in a leading role or supporting role, director, screenplay, cinematography, art direction, costume design, score, original song, and others, can play the role of properties and contribute to the Oscar for Best Movie. They focus on the 270 movies nominated between 1950 and 2003, and regress the value of the Best Movie Oscar (1 if the movie received this Oscar, 0 otherwise, that is if it was nominated but did not win) on the ratings of properties (which take the value 1 if the movie received the Oscar for this property, 0 otherwise). The results show that the important “properties” are actor in a leading role, director, screenplay, costume design and film editing; if these are awarded an Oscar, they contribute positively to the Best Movie award. Again, the weights are very different (0.668 for actor in a leading role, 2.342 for director, 1.157 for screenplay, 1.059 for costume design, and 0.614 for film editing). The regression equation can be used to rank the 270 movies during the 54 years, and is in disagreement with the Academy in nine years only out of 54. It can also “predict” the odds of receiving the award for best movie for out of sample years, using the other five awards, just before the one for Best Movie is announced at the very end of the celebration.

*Aggregating ratings or rankings of properties produced by a jury*

In most other cases, the difficulties are compounded, since one adds the problem of aggregating individual choices made by judges (and Arrow’s Impossibility Theorem), to the weighting problem (described by Dickie).

*Special cases*
Both difficulties can however be avoided if the properties that are taken into account are “objective,” and universally accepted. Economists draw a difference between what they call vertical and horizontal properties in goods. Consumers (or judges) can be unanimous in agreeing that more of a given vertical property say, the dimension of a flat, provides more utility or pleasure. But works of art are much better understood as being horizontally differentiated—a work by Rubens is preferred to a work by Beuys by some consumers, a Beuys to a Rubens by others, even if all objective properties (dimensions, for example) are identical.9

Therefore, in cases in which the jury agrees on some simple rules that are easy to implement and to verify, such difficulties vanish. This is so for racing competitions, jumping events, discus or javelin throwing, swimming, all cases where metering is objective, especially with today’s devices. The winner and the ranking are observable (all judges cannot but agree with the ranking), even if winning is the outcome of a complex process that involves many properties including intrinsic talent: A racer can run in an elegant way, a tennis player can be fair, and have beautiful backhand strokes, but these are not the qualities that are taken into account in the final judgment, which is objective and based on observable rules that cannot be infringed. This is also so for many games (chess, scrabble, backgammon, crossword puzzles), races in mathematical problem solving, but will not work for skating, diving or essay writing competitions.

9 See e.g. Gabszewicz and Thisse (1986) for a good exposition of the distinction between vertical and horizontal differentiation.
CONCLUDING REMARKS

Let us conclude by quoting again the great Gombrich (1979, p. 164):

“Early in November I read in the colour magazine of the Observer that Michelangelo is ‘out’. Those of us who believe in the objectivity of artistic values will be sorry to hear that, for if it were true the loss would be ours. He was not called great because he was famous. He was famous because he was great. Whether we like or dislike him, his greatness is an element in the story we are appointed to tell. It forms part of the logic of situations without which history would sink into chaos.”

These sentences summarize the main ideas of the paper: Artistic judgments are difficult, but the test of time makes them more “objective,” and helps sorting out those whose influence was determinant, and may be and have often been called geniuses.

But this does not tell us much about how to sort out contemporary geniuses, whom the media, stimulated by our competitive society (or should causality go in the other direction) keeps arousing. The 2,500 pages thick 2005 directory Awards, Honors & Prizes lists over 33,000 prizes covering some 400 subject headings including the arts. But most competitions fall short of discovering more than what consists of transitory fashion, and one can wonder whether English (2005, pp. 25 and 51) has not the last word in suggesting that, above all, they provide “a closed elitist forum where cultural insiders—artists, critics, functionaries, sponsors, publicists, journalists, consumers, kibitzers—engage in influence peddling and mutual back scratching.”

Bela Bartok is often quoted for having said that competitions are for horses, not for artists, though there exist horse races organized by some Mongol tribes, where the losing horse is rewarded and praised (Hayamon, 2012, p. 281).
References


Ginsburgh, Victor and Sheila Weyers (2012), Nominees, winners and losers, Manuscript.


Gombrich, Ernst and Quentin Bell (1976), Canons and values in the visual arts, a correspondence, *Critical Enquiry* 2, 395-410.

Graddy, Kathryn (2012), The extraordinary art critic Roger de Piles (1635-1709): An empirical analysis of his rankings and sale prices, Department of Economics, Brandeis University, manuscript.


Levinson, Jerrold (1990), Artworks and the future., in Jerome Levinson, *Music, Art and


Milo, Daniel (1986), Le phénix culturel: De la résurrection dans l'histoire de l'art, Revue Française de Sociologie, 27, 481-503.


Richardson, Jonathan (1719), An essay on the whole art of criticism as it relates to painting, London.


Simonton, Dean (2004), Film awards as indicators of cinematic activity and achievement: A quantitative comparison of the Oscars and six alternatives, Creativity Research Journal, 16, 163-172.


Verboord, Marc (2003), Classification of authors by literary prestige, *Poetics* 31, 259-281.


