
Natural Selection in the Expressional Principles in Darwin's *Expression of Emotions* (1872)

Hongjin Liu, PhD student in the University of Leeds
email: 632462362@qq.com

CSC* funded

Abstract: *Published in 1872, The Expression of the Emotions in Man and Animals is the only psychological book Darwin wrote. Based on the observations from him and others, Darwin concludes three principles to explain emotional expression of both human and animals. It is currently believed that the principles are non-Darwin because it is hard to find natural selection, the vital finding of Darwin's former works, in the book. This article argues the expression of emotions is regarded by Darwin as a kind of instinct and subject to natural selection. A statistical inquiry into the observational evidence in the book shows the differences among the three principles. Natural selection functions in the first two principles, the second principle is subordinate to the first. They also differentiate in scope of application and level of innovation.*

Key words: *natural selection, expressional principles, emotional expression, Darwin, statistics*

1. Introduction

The *Expression of Emotions* (1872) has scientifically opened up the study on emotional expression, since Darwin in the beginning proposes three expressional principles and lists six research methods. In the first three chapters of the book, Darwin maintains the core findings of his research on emotional expression: the three general principles governing emotional expression of both human being and animals. They are: Principle 1 'Serviceable Associated Habits'; Principle 2 'Antithesis'; Principle 3 'direct action of the nervous system' (Darwin, 1872, p. 28-29). The *Expression of*

* The project is funded by China Scholarship council (2012-2015).

Emotions is regarded by Montgomery (1985) as non-Darwinian compared to the *Origin of Species* (1859), due to his failure in looking for the vital finding in Darwin's prior work— natural selection. Montgomery believes Darwin adopts Lamarckian use-inheritance, rather than natural selection, as the evolution mechanism of emotional expression. However, he does not see its relation to the prior book of Darwin, the *Descent of Man* (1871).

In Chapter II and III of the *Descent of Man* (1871) regarding mental capacity, Darwin faces difficulty to verify that the mental power of human beings originates from the capacity of lower animals. So in the *Expression of Emotions* (1872) he diverts the attention to focus on the outer behavior— 'expression' rather than the inner capacity— 'emotion'. Such diversion not only helps Darwin to better defend his theory on evolution, but also scientifically establishes the study of expression (of emotions).

Furthermore, Montgomery and many other scholars have long been taking as granted that the three expressional principles are believed by Darwin as pertinent to both human and animals without differentiations: no difference has been suggested in their scope of application. It is because there lacks an inspection into the data provided in the relevant chapters by Darwin. This article mentions this problem with a statistic on the evidence from both human and animals in the *Expression of Emotions* to demonstrate such distinctions. The first three sections below regard the three principles respectively with an introduction to their theoretical basis and observational evidence. Under the title of each of the sections, Darwin's implications of the principles are given in italics with the perspective of instinct.

2. First principle: Serviceable Associated Habits

Some of the emotional expressions, as kinds of instincts, are formed by 'Serviceable Associated Habits' and developed by 'natural selection'.

Definition in the book

Movements (motions) under certain emotion had a tendency to release the emotion, the association between the motions and emotion was externally observed as 'habits', which could be inherited by the next generation (Darwin 1872:28).

2.1 Origin of the principle

The embryonic form of the first principle has already been recorded in Darwin's 1838 Notebook M: *Metaphysics on Morals and Speculations on Expression*. It is noted on p46, where the word 'association' first appears, that 'when a muscle is moved very often, the motion becomes habitual & involuntary..... An intentionally recollection of anything is solely by association, & association is probably a physical effect of brain the similar remark thoughts, being functions of same part of brain, or the tendency to habit of producing a train of thought' (Darwin 1838: 46).

The above supposition called to Darwin when considering the formation and inheritance of instinct, that is, the observable 'habits' originally applied to instinct. Later this principle is employed in the study of emotional expression, since Darwin believes some intricate expressions are kinds of instinct for they burst out unconsciously under certain emotions and could be inherited by the offspring. In Chapter I relating the first principle, Darwin devotes several pages (p35- p42) to talk about the relationship between reflex action (instinct) and emotions. He argues 'some reflex actions can hardly be distinguished from instincts; and, of the latter, it may be added, some cannot be distinguished from inherited habits' (Darwin 1872: 35). His further argument implied the correlation between Principle 1 and his prior finding— natural selection: 'although some instincts have been developed simply through long-continued and inherited habit, other highly complex ones have been developed through the preservation of variations of pre-existing instincts— that is, through natural selection' (p. 41). On the ground of the above narrations, I conclude the hinted idea of Darwin as shown in the above italics.

2.2 Evidence given by Darwin

After stating the first principle, Darwin in Chapter I has to provide some evidence to uphold and illustrate it. He in all gave 48* examples of human and animals' expression of emotions. Among them, 31 cases are from animals and 17 examples are from human beings. In the cases of animals, the one appears most times is dog, eight times, the secondary is horse, four times, the third are cat and jackal, both three times, then are cattle, wolf, fennec, pig and six kinds of birds, each of them appears once. The lowest animals (amphibian and insects) are frog, a kind of moth† and caterpillar, all appears once. In the cases of human being, three of them concern infants while the other 14 are about adults. One thing worthy mention is that both the cases of infants and the three lowest animals refer to instincts‡, while the others all pertain to the expression of emotions.

2.3 Conclusion

From the above testimony it can be seen that the Principle of Serviceable Associated Habits is illustrated by Darwin to govern the formation and inheritance of both instinct and expression of emotions. He consciously mixes up, or say treats equally, the reflex action, instinct and the expression of emotions. In fact, in Chapter I no single word is about the differences among the three, but inversely, all he writes is to convince the readers the sameness between the expression of emotion and instinct both by logical reasoning (see Section 2.1) and by promiscuous exemplification (see Section 2.2). Through this way could Darwin link the higher animals (mammal including human being) and the lower animals. That is, by persuading the readers the expression of emotions is a kind of instinct, could he verify those species are equal, or at least that they follow the same law— natural selection functions in the

* The number of cases is counted in the way that when several animals are exemplified for one expression of emotions, it is counted as once and attached to the less appeared animals. For example, when Darwin says dog, horse and pig express startle, it is counted as 'pig once' for the other two animals have appeared many times in the case of other expressions of emotions.

† Humming-bird Sphinx-moth (*Macroglossa*) (p. 30).

‡ The amphibian and insects certainly have no observable expression of emotions, while the cases on infants are all about reflex actions, such as sneer.

formation and inheritance of instinct in lower animals and expression of emotions in higher animals (see the cases given in Section 2.2).

Till now we can see that this principle is reasonable to come first because it conveys the proposition expressed in Darwin's prior research: the *Origin of Species* (1859) and the *Descent of Man** (1871). The merits of the three books are coherent that in the 1859 and 1871 books Darwin verifies that the evolution of animals and human's external characters and mental capacity follow the same law: natural selection, and then in the *Expression of Emotions* he further states that the same law also governs the field of the expression of emotions in man and animals. In the *Descent of Man*, Darwin has tried to verify that human and animals' emotion both derive from social instincts but his verification is limited due to lack of evidence. So, in the *Expression of Emotions*, merely does he focus on the expression of emotions—the outer manifestation of mental state. Thus the logic chain of all his works is made coherent and intact.

3. Second principle: Antithesis

Antithesis expression of antithesis emotion is an instinct

Definition in the book

'When a directly opposite state of mind is induced, there is a strong and involuntary tendency to the performance of movements of a directly opposite nature, though these are of no use; and such movements are in some cases highly expressive' (Darwin 1872:50).

In the definition quoted ahead, it is stated Darwin's logic sequence: 'the movements of a directly opposite nature' are wider than expressive movements that they are only 'in some cases highly expressive'. In other words, the Principle of Antithesis governs broader movements including the expression of emotions. It is indeed another finding in

* The *Expression of Emotions* is at first designed to be one chapter of the *Descent of Man* (1871), but its content has largely exceeded the volume, so is separately published in the next year as the book we can see nowadays (Darwin 1958: 131; Prodder 2009: 6).

Darwin's thinking of instinct and is the most original one as shown in Darwin's evidence provided below.

3.1 Observational evidence

Unlike the first one, the second principle does not appear either in Darwin's notebooks* or in the *Descent of Man*. Darwin obtains it from his own observation without citation from other scholars†. Three observational examples are given in Chapter II of the *Expression of Emotions* to illustrate it: two on dogs and one on cats. The first two examples are generally descriptions that when a dog/cat is happy or attempts to show kindness to its host, its behavior is exactly opposite to the status when it is angry or offensive (Darwin 1872:64). Apart from the non-specific description, Darwin gives another particular case of his own dog. Its pleasure to go out walking and disappointment when stops walking are described in detail by Darwin as the 'instantaneous and complete change of expression', an instance of antithesis expression of antithesis emotion.

3.2 Explanation of Darwin

In talking about the first principle, Serviceable Associated Habits, Darwin has argued that if an expression is serviceable, then it can be reserved and inherited by the process of natural selection. In order to employ natural selection, as he deal with the first principle, to explain the arising and heredity of the antithesis expression, Darwin has to verify that the antithetical movements are 'serviceable', or say beneficial to animals and human being. The only benefit of the antithesis expression of emotions comes up to Darwin is for communication. Thus Darwin this time admits that the discussion of this principle is limited in social animals (or animals with social instincts‡).

* As Montgomery points out, on p146-147 of the Notebook M, Darwin notes an example of antithesis expression of dog (Montgomery 1985:41), but it at that time exists as an observational fact rather than a general principle. The Principle of Antithesis should come up after Darwin has compiled his files.

† Except for the several drawings used to illustrate the antithesis expression of emotions.

‡ The words in () are used in the *Descent of Man*. In fact, the first two principles are legacies of, or at least say coherent from the thoughts in Chapter II and III of the 1871 book. In these two chapters, *Comparison of the mental capacity of man and the lower*

‘With social animals, the power of intercommunication between the members of the same community,—and with other species.....,—is of the highest importance to them’ (Darwin 1872:60). He says social animals have their own language and some kinds of them could even understand human language and gestures. This principle is essential to invent their language, including conventional signs and sign language (Darwin 1872:61-62).

3.3 Comparison and discussion

The second principle has aroused many criticisms*, which is due to some problems of both Darwin’s strategy and the readers’ understanding on it. I now consider it with the first principle to reveal their relationship and the problems around it.

Firstly, Darwin has stated, but not highlight, that the second principle is subordinate to the first one: it is a particular case subject to the function of the first principle. In the last paragraph of Chapter II talking about the second principle, Darwin says: ‘As the performance of ordinary movements of an opposite kind, under opposite impulses of the will, has become habitual in us and in the lower animals, so when actions of one kind have become firmly associated with any sensation or emotion, it appears natural that actions of a directly opposite kind.....should be unconsciously performed through habit and association, under the influence of a directly opposite sensation or emotion’ (Darwin 1872:65). In short, since the antithesis expression is serviceable (see Darwin’s explanation quoted in Section 3.2 above), then it ‘should be unconsciously performed through HABIT and ASSOCIATION’ under antithesis emotion. Thus we can see that, the second principle is not an independently governing principle but a

animals and —*continued*, it is obscure in understanding the ‘social animals’ there for Darwin tends to recognize all animals as bearing social instincts— a clear classification has never been made either in the *Descent of Man* or in the *Expression of Emotions*.

* See Sully, 1874, *Sensation and Intuition*, p. 29; Mantegazza, 1885, *La Physionomie*, p. 76); and Dumont, 1877 *Théorie Scientifique de la Sensibilité*, 2nd edit., p. 236.

particular case under the effect of the first principle, Serviceable Associated Habits.

Secondly, Darwin has stressed in the last sentence of Chapter II that 'there cannot be a doubt that several expressive movements due to the principle of antithesis are inherited' (Darwin 1872:65). It is an obscure narration if the readers do not link it with Darwin's study on instinct either in the *Descent of Man* or in his earlier thoughts noted in the Notebook M. In fact, as I argued in the beginning, these principles are results of Darwin's thoughts on instincts. Especially to the first one, Darwin aims to verify that the instinctive expression is inheritable through the effect of natural selection (see Section 1). Further in talking about Principle 2, he attempts to testify that the antithesis expression is another case of instinctive expression and is also inheritable under natural selection. In the *Expression of Emotions*, Darwin is too aggressive to raise the second principle as a separate and general principle, which is indeed a particular case governed by the first one. And if the readers do not bear in mind Darwin's emphasis on the heredity and natural selection, they might criticize the second one due to its unsuitable status aggressively raised by Darwin.

4. Third principle: Direct action of the nervous system

Physiological reaction is a kind of instinct

Definition in the book

'Certain actions, which we recognize as expressive of certain states of the mind, are the direct result of the constitution of the nervous system, and have been from the first independent of the will, and, to a large extent, of habit' (Darwin 1872:66).

This principle originates from both the observational evidence Darwin collects and a theoretical summary of his contemporary philosophers and physiologists. The following two sections will introduce these two sources respectively.

4.1 Observational evidence

From his own and others' observation, Darwin in all gives seven groups of expressions, which he believes are subject to the third principle. They are: trembling of the muscles, heart beating, perspiration, respiration, modified secretions, unintentional sound and change of color in the hair. These expressions on the list are sequenced by their amount of examples provided in the book*. The first six of them are exemplified with cases from both human and animals, while the last one only has one human example cited from other scholars. That is, no animal is observed losing color of their hair or feather. The lowest animals exemplified are two kinds of birds with respect to trembling of the muscles and loss in skin color.

4.2 Theoretical citation

The definition given in the beginning of the section shows completely the implication of the third principle: some expressions are independent from the first two principles and induced directly by nervous system— a kind of instinct as written in the italics. What I want to stress here is the first half— independence from the first two principles— is actually the core text of this principle, as shown in the first example raised by Darwin in the book†. On the contrary, the other half about nervous system is much less innovative, since it is cited, without paraphrasing, from the physiologists, such as Müller and Spencer: it serves to explain the observational evidence as summarized above—the seven groups of expressions. For example, to interpret the trembling of muscles, Darwin cites from Müller's *Elements of Physiology* (1838) that 'any strong excitement of the nervous system interrupts the steady flow of nerve-force to the muscles'. Thus we can see that Principle 3 is built on the observational evidence discussed in Section 4.1, while

* Though of the least example, the last expression on the list— change of color in the hair— is introduced firstly in Chapter III of the *Expression of Emotions* concerning the third principle. It is because in Darwin's mind, it is the best case to demonstrate the effect of nervous system which is 'independent of the will, and, ... of habit', or in other words independent of the effect of the first principle.

† Same as above footnote.

his citation and discussion on nervous system merely serves to explain those observational facts.

5. Concluding discussion of the three principles

5.1 Their timing and importance sequence

As mentioned above, the first Principle has already been noted in Darwin's Notebook M in 1838 and also appears in the *Descent of Man* (1871) concerning the heredity of instincts. It contains many legacies from the author's former research, such as the theory of natural selection and his thoughts on the animal's instincts. Among the three principles, without no doubt that it is the first one obtained by Darwin. In regard to the other two principles, there is also no doubt that they arise after a certain amount of observational facts have been collected and compiled. The thinking process of Darwin on the topic of emotional expression would be like this:

Darwin classifies these facts and finds the Serviceable Associated Habits (or inheritable instincts under the effect of natural selection*), which has arrived to him much earlier, also governs the arising and heredity of most expressive movements he observes or collects in man and animals. So he gives out the first principle as the basic and governing one. However he also finds some expressions of emotions unserviceable/useless to animals (and human being). Then he tries to explain the existence and inheritance of these expressions. They are divided into two kinds: the antithetical expression pairs and the seven groups of expressions which he believes are induced by nervous system. Then the other two principles are brought out to label and define those expressions. After that, Darwin has to provide theoretical ground for those principles to account for their existence and inheritance. Darwin attempts to link them with the first principle, which, in his mind, has been proved as ground. If he could verify those expressions as useful, then the core text of the first principle, natural selection, could also govern the heredity of those expressions. His trial with the second principle is successful, at least in his own arguments. The

* See the italics under the title of Section 1.

antithetical motions are claimed as essential to invent human and animals' languages (conventional and gestural) which are useful for communication. Thus, the natural selection, the invariable theme of Darwin from 1859, functions again in the heredity of the antithesis expressions: since useful, then are selected and inherited. That is why the Principle of Antithesis, as I argued in Section 2, is subordinate to the first principle: they are serviceable, so are subject to the Serviceable Associated Habits.

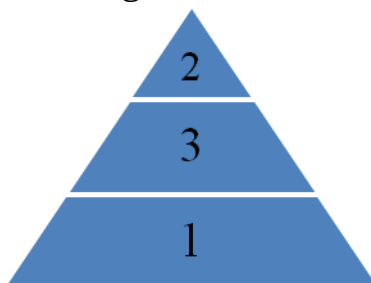
However, with the third principle, Darwin fails to see any usefulness in some of the expressions involved, such as the loss of color in hair. When his handy weapon— natural selection— loses its effect, this principle could merely serve to describe and label the certain kind of expression of emotions.

By now we can see that the theoretical ground of the first principle came to Darwin earlier than his gathering the observational facts, while the other two principles are obtained after the classification of those facts. In Darwin's mind, the first principle plays a fundamental and basic role to explain most of the expressive movements. The other principles are firstly used to categorize certain kinds of expression of emotions. That is, they are descriptive and explanatory principles without governing effect as the first one.

5.2 Their innovation degree

The 'innovation degree' here means their independence from Darwin's prior works and other scholars.

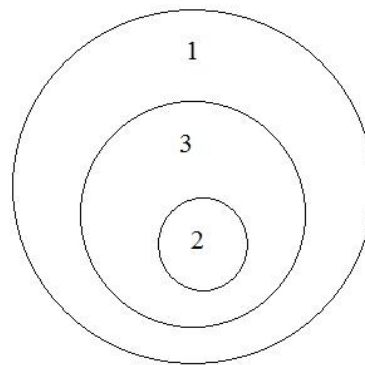
The innovation degree of the three principles



As shown in the above pyramid, the first principle relates most to Darwin's prior research, so bears the lowest degree of innovation in the study of emotional expression. Its theoretical base was previously used in Darwin's former works. As mentioned above, the other two principles are invented to label and define certain kinds of expressive movements. To categorize them contributes originally to the relevant research. Between the two principles, Principle 2 is regarded more innovative than the third. After drawing them from the observational facts he collected, Darwin sorts for their academic base. To the third principle, all of its theoretical sources (of nervous system) are cited from physiologists. With the Principle of Antithesis, Darwin states originally the mechanism of antithesis expressions without any citations from other scholars.

4.3 Their application scope

Scope of the animals involved



From the statistic of the examples provided in the *Expression of Emotions* (see the sections ahead titled 'evidence': 2.2, 3.1, 4.1), I draw the above figure to demonstrate the scope of the animals (including human) involved. As shown in the area of these circles, the animals exemplified for Principle 1 are of the broadest range. They

include from the highest mammal to the animals as low as six kinds of birds, frog, moth and caterpillar, while the third principle is only applicable to those with nervous system. Though Darwin does not make clear what kinds of animals have nervous system, the lowest animals appear are two kinds of birds. With the smallest range, the third principle is only exhibited in advanced mammals: human, dog and cat.

6. References

- Darwin, Charles. (1838). *Notebook M: [Metaphysics on morals and speculations on expression]*. Transcribed by Kees Rookmaaker, edited by Paul Barrett. (Darwin Online, <http://darwin-online.org.uk/>)
- Darwin, Charles. (1871). *The Descent of Man, and Selection in Relation to Sex*, vol.1. London, England: John Murray.
- Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. London, England: John Murray.
- Darwin, C. (1958). *The Autobiography of Charles Darwin, 1809-1882, with Original Omissions Restored. Edited with Appendix and Notes By Nora Barlow*. Collins.
- Montgomery, William. (1985). Charles Darwin's thought on expressive mechanisms in evolution, in: G. Zivin (Ed.), *The development of expressive behavior. Biology-environment interactions*, Academic Press, p.27-50.
- Prodger, Phillip. (2009). *Darwin's camera: Art and photography in the theory of evolution*. Oxford University Press.