

Name of the author: Haengjin Choe

E-mail address of the corresponding author: eupjeong@gmail.com

**Keywords:** psychophysics; Weber–Fechner law; information theory; Shannon entropy; response information; perception entropy; percentropy

## **Abstract**

Since the publication of Shannon's article about information theory, there have been many attempts to apply information theory to the field of neuroscience. Meanwhile, the Weber–Fechner law of psychophysics states that the magnitude of a subjective sensation of a person increases in proportion to the logarithm of the intensity of the external physical-stimulus. It is not surprising that we assign the amount of information to the response in the Weber-Fechner law. But no one has succeeded in applying information theory directly to that law: the direct links between information theory and that response in the Weber-Fechner law have not yet been found. The proposed theory unveils a link between information theory and that response, and differs subtly from the field such as neural coding that involves complicated calculations and models. Because our theory targets the Weber–Fechner law which is a macroscopic phenomenon, this theory does not involve complicated calculations. Our theory is expected to mark a new era in the fields of sensory perception research. Our theory must be studied in parallel with the fields of microscopic scale such as neural coding. This article ultimately aim to provide the fundamental concepts and their applications so that a new field of research on stimuli and responses can be created.

### 1.1 Psychophysics of Weber and Fechner

In 1834, the German physiologist Ernst Heinrich Weber carried out an exciting experiment. Weber gradually increased the weight which the blindfolded man was holding. With doing so, Weber asked the subject when he first felt the increase of weight. Through this experiment, Weber found out that the least perceptible difference in weight was proportional to the starting value of the weight. For instance, if someone barely feels the difference between 100 g and 104 g, then they barely feel the difference between 150 g and 156 g. In this instance, 100: 104 = 150: 156 holds. Weber's findings were elaborated in his book (1851).

Weber's empirical observations were expressed mathematically by the German physicist Gustav Theodor Fechner who was a Weber's student (Fechner 1860). Fechner called his formulation Weber's law. Fechner applied the result of Weber's experiment to the measurement of sensation (Ward 1876; Houstoun and Shearer 1930). The Weber–Fechner law established in the field of psychophysics attempts to mathematically describe the relationship between the magnitudes of external physical-stimulus and the perceived intensities of the stimulus (Fechner 1860). In this fascinating law, the physical stimuli are weight, sound frequency, etc. The Weber–Fechner law states that the relationship between stimulus and its perception is logarithmic. Therefore, if a stimulus varies as a geometric progression,

the corresponding perception varies as an arithmetic progression. Bear in mind that the stimuli in the Weber–Fechner law have physical units but that the responses in that law have no physical units. The best example of the Weber–Fechner law is that many musicians adopt the equal temperament when they tune their musical instruments. Since the nineteenth century, musicians have tuned virtually all instruments in equal temperament (Berg and Stork 2005). Equal temperament in music theory is a tuning system in which the octave is divided into 12 semitones of equal size; note here that the size does not indicate frequency, but musical interval. In the equal temperament system, the musical instruments are tuned so that the frequencies of all notes form a geometric sequence. Such a tuning results in the fact that the musical intervals of every pair of two consecutive notes are the same (Hopkin 1996; Berg and Stork 2005). This example needs to be explained in more detail. Transposing or transposition in music means changing the key of a piece of music. When the musical-instrument accompaniment is too high or too low for us to sing to, we change the musical accompaniment into a lower or higher key. Even if we transpose the accompaniment into another key, as if by magic, we don't feel awkward. There is an essential reason that the accompaniment does not sound awkward. This is because the musical instruments were tuned to the equal temperaments. Such a tuning gives an equal perceived step size because pitch is perceived as the logarithm of frequency. As another concrete example of the Weber-Fechner law, the response of the human ear to change in sound intensity is logarithmic. Therefore, it is convenient to speak of the sound level defined as

above the standard threshold of hearing intensity (Kinsler et al. 2000). The standard threshold of hearing intensity is chosen as  $10^{-12}$  W/m<sup>2</sup>, which is near the lowest limit of the human range of hearing (Kinsler et al. 2000; Halliday et al. 2014). Note here that because the defining equation of the decibel scale has the same form as the Weber–Fechner formula, the perceived magnitude is proportional naturally to a number obtained after we drop decibel in the sound level. As a third concrete example of the Weber–Fechner law, the response of the human eye to brightness of light is logarithmic. The stellar magnitudes  $m_1$  and  $m_2$  for two stars are related to the corresponding brightnesses  $b_1$  and  $b_2$  through the equation

$$m_2 - m_1 = \sqrt[5]{100} \log \frac{b_1}{b_2},$$

which results from the Pogson's proposal that a difference of five magnitudes should be exactly defined as a brightness ratio of 100 to 1 (Pogson 1856; Burke-Gaffney 1963; Mayer 1986; Young 1990; Hearnshaw 1992; Schulman and Cox 1997). Note here that the stellar magnitude corresponds to the response and hence has no physical unit.

## 1.2 Information theory of Shannon

Claude Elwood Shannon defined mathematically the concept of 'the amount of information' which is sometimes called 'the information content' (Shannon 1948).

Going one step further, he defined mathematically the concept of entropy (Shannon 1948). Due to information theory, we came to be able to quantify the amount of information transmitted over a communication channel.

### 1.3 Background of digitization theory

Is it possible to digitize the Weber–Fechner law? Digitizing the Weber– Fechner law in this article means the following: we define mathematically a new kind of the information content from the Weber-Fechner law, and going one step further we define mathematically a new kind of entropy from that law. This digitization will be able to be possible if we discover links between the information content formula of Shannon and the Weber–Fechner formula. The author proposes a theory that we are able to digitize the Weber-Fechner law. The digitization procedure is as follows. We will add one more stimulus to the existing various stimuli applied to the Weber-Fechner law. This addition procedure is based on the structural analogy between the Weber-Fechner formula and the information content formula. This additional stimulus and the corresponding perception are extremely extraordinary. We will digitize the Weber–Fechner law by extending and reinterpreting that law. Digitization of the Weber–Fechner law which is achieved by merging information theory with the Weber–Fechner law will lead us to two new concepts. Firstly, we define a new concept which the author calls 'the amount of response information' or 'the response information content'. Ultimately, we end up defining another new concept which the author calls the 'perception entropy'.

## 2 Weber–Fechner law and the information content

### 2.1 Derivation of Weber–Fechner law

The Weber–Fechner law is derived as follows (Nutting 1908). In Weber's experiment, let S be weight at some instant and let dS be the differential increment in weight. Let dR be the differential change in perception. The equation to express the result of Weber's experiment is

$$dR = k \frac{dS}{S},$$

where k is the constant of proportionality. (There is one significant thing about k, which will be mentioned in Section 4.) Integrating both sides of this equation yields

$$R = k \ln S + C$$

in which case C is an integration constant. Let us suppose that  $S_0$  is the threshold of the stimulus below which a person does not perceive anything. Because R=0 whenever  $S=S_0$ , C must be equal to  $-k \ln S_0$ . Therefore, we can finally obtain

$$R = k \ln \frac{S}{S_0} = K \log_2 \frac{S}{S_0},$$

which is the mathematically expressed Weber–Fechner law. Let us note that R has no physical unit. But the physical stimulus S has, of course, a physical unit.

### 2.2 Definition of the information content

Shannon defined the amount of information of an event of the occurrence probability P as

$$I = -\log_2 P$$
,

where the information content I is measured in bits (Shannon 1948). And he defined the entropy of a discrete random variate as the average amount of information of the random variate (Shannon 1948). In order to derive a new kind of the information content and hence a new kind of the entropy, let's redescribe the above equation as

$$I = -\log_2 \frac{P}{P_0} = \log_2 \frac{P_0}{P},$$

where as a matter of course  $P_0 = 1$ .

# 3 Derivation and interpretation of digitization theory

## 3.1 Special derivation

In the equation

$$R = K \log_2 \frac{S}{S_0},$$

for all practical purposes  $S \ge S_0$ . In the equation

$$I = \log_2 \frac{P_0}{P},$$

for all practical purposes  $P \le P_0$ . We can interpret  $P_0 = 1$  as the threshold of P. Most importantly, we can interpret I as a response to P; it would be nice for us to call P a bizarre kind of stimulus. However, this bizarre stimulus fundamentally differs from other ordinary physical stimuli in that it is simply a number which has no physical unit. Very important another fact is that the higher P is, the lower I is while the higher S is, the higher R is. But it is not essential that I is a strictly decreasing function with respect to P while R is a strictly increasing function with respect to S. The essentially important facts are that R and I with respect to S and P, respectively, are logarithmic and that both S and P have the threshold values. From these viewpoints, consequently, the author cannot choose but say that the occurrence probability is an anomalous kind of stimulus. It would be appropriate that we name the occurrence probability the 'mathematical stimulus' in that the occurrence probability has no physical unit. This mathematical stimulus is under the government of the Weber–Fechner law. By widening the range of applications of the Weber– Fechner law, we come to be capable of combining information theory with the Weber–Fechner law of psychophysics.

For simplicity, in current subsection we set K to 1 for all sensations. (We make complete descriptions in Subsection 3.2. In that subsection we make no assumptions about K.) Then, we can say that the amount of perception is 0 bits when  $S = S_0$  and that the amount of perception is 1 bits when  $S = 2S_0$ . But there is one thing to be careful about. In reality, different people have different  $S_0$ s. So, different

people have different amounts of response information for the identical S. In addition, as the body of a person ages,  $S_0$  is changed. So, for the identical S the response information content of even a person changes with age. Because of this fact, the response information content is quite different from the information content of Shannon. In the defining equation of the information content of Shannon,  $P_0$  is fixed at 1.

If the physical stimulus is the weight, the problem is simple. In order to clearly define a new kind of entropy, let's consider the sound wave from now on. Because a pure tone which is a sound wave with a sinusoidal waveform is characterized by both a frequency and a displacement amplitude, considering only either frequency or amplitude may be meaningless. Thus, we consider both simultaneously. For the sake of argument let's consider a pure tone of frequency f and amplitude f. Physics entropy and Shannon entropy are subject to the constraint that the total sum of probabilities is 1. On the other hand, there is not such a constraint in the defining equation of the new kind of entropy. The physical quantities f and f change much more freely. Therefore, it is impossible that we define the new entropy in the same form as the Shannon entropy. In order words, we must not define the new entropy as

$$\frac{A}{A_0}\log_2\frac{A}{A_0} + \frac{f}{f_0}\log_2\frac{f}{f_0},$$

which is of the form of the Shannon entropy. Before we define the new entropy, we first define the total amount of response as

$$\log_2 \frac{A}{A_0} + \log_2 \frac{f}{f_0} = \log_2 \frac{Af}{A_0 f_0},$$

where  $\log_2 \frac{Af}{A_0 f_0}$  is measured in bits. In other words, the total amount of response is defined so that it equals the sum of the amounts of each individual response. Consequently, we define the new entropy as the arithmetic mean of the response information contents being considered; that is, new entropy is defined as

$$\left(\log_2 \frac{Af}{A_0 f_0}\right)/2.$$

The unit of this quantity is bits/response. Please note that in the current situation there are two responses: one response due to the amplitude and the other due to the frequency. The author calls this new quantity the 'perception entropy' or the 'percentropy'. The term 'percentropy' is a new compound word formed from perception and entropy. The percentropy does not correspond perfectly to the Shannon entropy. Nonetheless, it is useful to define the concept of percentropy. One of reasons is that the percentropy is linked directly with the physics energy. We are easily able to understand such a fact. As an example, consider two pure tones. One tone has amplitude  $A_1$  and frequency  $f_1$ . The other tone has amplitude  $A_2$  and frequency  $f_2$ . The percentropy of the former tone is  $\left(\log_2 \frac{A_1 f_1}{A_0 f_0}\right)/2$ . The percentropy of

the latter tone is  $\left(\log_2 \frac{A_2 f_2}{A_0 f_0}\right)/2$ . The  $A_1 f_1 = A_2 f_2$  is a necessary and sufficient condition for the fact that the percentropies of the two tones are the same. And,  $A_1 f_1 = A_2 f_2$  if and only if  $(A_1 f_1)^2 = (A_2 f_2)^2$ . We can therefore immediately conclude that the  $A_1 f_1 = A_2 f_2$  is a necessary and sufficient condition for the fact that the energies of the two tones are the same. Let us now consider two compound tones. One compound tone consists of a pure tone which has amplitude  $A_{11}$  and frequency  $f_{11}$  and a pure tone which has amplitude  $A_{12}$  and frequency  $f_{12}$ . The other compound tone consists of a pure tone which has amplitude  $A_{21}$  and frequency  $f_{21}$  and a pure tone which has amplitude  $f_{22}$  and frequency  $f_{23}$ . The percentropy of the former compound tone is

$$\left(\log_2 \frac{A_{11}}{A_0} + \log_2 \frac{f_{11}}{f_0} + \log_2 \frac{A_{12}}{A_0} + \log_2 \frac{f_{12}}{f_0}\right) / 4 = \left(\log_2 \frac{A_{11}f_{11}A_{12}f_{12}}{{A_0}^2 {f_0}^2}\right) / 4.$$

The percentropy of the latter compound tone is

$$\left(\log_2 \frac{A_{21}}{A_0} + \log_2 \frac{f_{21}}{f_0} + \log_2 \frac{A_{22}}{A_0} + \log_2 \frac{f_{22}}{f_0}\right) / 4 = \left(\log_2 \frac{A_{21}f_{21}A_{22}f_{22}}{A_0^2 f_0^2}\right) / 4.$$

The  $A_{11}f_{11}A_{12}f_{12} = A_{21}f_{21}A_{22}f_{22}$  is a necessary and sufficient condition for the fact that the percentropies of the two compound tones are the same. And,  $A_{11}f_{11}A_{12}f_{12} = A_{21}f_{21}A_{22}f_{22}$  if and only if  $(A_{11}f_{11})^2(A_{12}f_{12})^2 = (A_{21}f_{21})^2(A_{22}f_{22})^2$ . Therefore, in this compound-tones situation, the equality of the two percentropies is a necessary

and sufficient condition for the fact that the products of energies of the two pure tones composing each compound tone are the same.

#### 3.2 General derivation

Now we do not assume that *K*s are 1 for all sensations. Because *K*s are simply constants, we just have to define the response information content for a single stimulus as

$$K \log_2 \frac{S}{S_0}$$

where  $K \log_2 \frac{S}{S_0}$  is measured in bits. Let's consider a pure tone of frequency f and amplitude A. Let  $K_A$  and  $K_f$  be the corresponding constants. We first define the total amount of response as

$$K_A \log_2 \frac{A}{A_0} + K_f \log_2 \frac{f}{f_0} = \log_2 \left(\frac{A}{A_0}\right)^{K_A} \left(\frac{f}{f_0}\right)^{K_f},$$

where  $\log_2\left(\frac{A}{A_0}\right)^{K_A}\left(\frac{f}{f_0}\right)^{K_f}$  is measured in bits. In other words, the total amount of response is defined so that it equals the sum of the amounts of each individual response. And then we define the new entropy as the arithmetic mean of the response information contents being considered; that is, new entropy is defined as

$$\left\{\log_2\left(\frac{A}{A_0}\right)^{K_A}\left(\frac{f}{f_0}\right)^{K_f}\right\}/2.$$

The unit of this quantity is bits/response. Because the amplitude and frequency change independently of each other, we took the arithmetic mean. Notice that the percentropy is not linked directly with the physics energy unless  $K_A$  equals  $K_f$ .

## 3.3 Utilization and interpretation

According to our theory, it comes to be possible that we utilize percentropy (or the total amount of response information) for quantifying the sensitivities of the sensory organs of humans; at this time the sensitivity is a function of stimuli and their thresholds. This quantification allows us to compare the sensitivities of the sensory organs of people. Amazingly, the sensitivity of the sensory organ implies the superiority of the sensory organ. It is best to explain an example. For simplicity, we set K=1 for all sensations. Let us consider the case of the hearing. There are two people. For a person K=1, K=1,

$$\left(\log_2 \frac{Af}{A_0 f_0}\right)/2 = \left(\log_2 \frac{330 \times 1000}{33 \times 20}\right)/2 \approx 4.483 \text{ bits/response,}$$

and that of Y is

$$\left(\log_2 \frac{Af}{A_0 f_0}\right)/2 = \left(\log_2 \frac{330 \times 1000}{22 \times 60}\right)/2 \approx 3.983 \text{ bits/response.}$$

Thus the hearing organ of X is superior to that of Y. In fact, as we can very easily understand, we only need to know each individual's  $A_0f_0$  in comparing superiority of the hearing organs of people. Consequently, we can say that the information-theoretical measure of the superiority of the hearing organs of people is  $A_0f_0$ . Of course, we might effortlessly devise  $A_0f_0$  as a measure of the superiority, without any knowledge appearing in this article. The key point here is that the measure  $A_0f_0$  is justified by the digitization theory of the Weber–Fechner law; i.e.,  $A_0f_0$  is valid from the information-theoretical perspective. Please note that we have set K=1 regardless of the type of sensation.

Even if we do not assume that K = 1 for all sensations, we can still compare the sensitivity of the sensory organs of people. But the measure of the superiority is no longer  $A_0 f_0$ . If we assume that the values of  $K_A$ s for people are all identical and that the values of  $K_f$ s for people are all identical, then the measure of the superiority is  $A_0^{K_A} f_0^{K_f}$ .

So these are good. The question is then what 4.483 bits/response and 3.983 bits/response mean; equivalently, what 8.966 bits and 7.966 bits mean. Suppose constant physical-stimuli are exerted. It is clear: that the response information content is proportional to the duration of the response; that there is no significant difference between the duration of the response and the duration of the external stimulus. And, probably, the duration of the response is quantized. In order words, the duration of the response is always some integral multiple of elementary

duration. According to this hypothesis, the response information content during some time interval equals some integer times the response information content during elementary duration. Consequently, whenever talking about the numerical value of K, we must mention the duration of the response together. But it is cumbersome to do so. Therefore, it is convenient to determine the numerical value of K for elementary duration first, and then use the value so determined. Now, we can say what 8.966 bits and 7.966 bits mean. The 7.966 bits is a certain number times the response information content for elementary duration.

# 4 Future challenges

The stimuli and responses that have been mentioned so far are about the external physical-stimuli and their perceptions; the stimuli and responses that have been mentioned so far do not include the stimuli and responses that neural coding handles. In order for the digitization theory of the Weber–Fechner law to be firmly established, we need to elucidate the relationships between microscopic and macroscopic theories on stimuli and responses. Past information-theoretical researches on stimuli and responses have mainly focused on microanalysis (MacKay and McCulloch 1952; Rapoport and Horvath 1960; Werner and Mountcastle 1965; Stein 1967; Borst and Theunissen 1999; Dimitrov and Miller 2001; Koechlin and Summerfield 2007; Ikeda and Manton 2009). Those researches involve complicated models and calculations. Because the Weber–Fechner law is a macroscopic phenomenon which is a manifestation of numerous microscopic outcomes, the

digitization theory of the Weber–Fechner law involves far less complicated calculations. In the future, macroanalyses as well as microanalyses should be done together.

There is one significant thing about K in particular. In psychophysics K is originally an experimentally determined constant. But, the numerical value of K for each sense in our theory must be determined by comparing our theory with works of microscopic scale.

Meanwhile, author guesses that regardless of the type of sense, a theoretic expression (not a measured value) for *K* includes Boltzmann's constant or the Boltzmann-like constant which serves to bridge the macroworld and the microworld. Once the *K*s and elementary durations are all determined, the field of research on stimuli and responses is expected to develop rapidly.

## **5** Conclusions

So as to achieve our goal, we must find links between the information content formula and the Weber–Fechner formula. These links are the key idea behind the digitization theory of the Weber–Fechner law. The link presented in this article is an undiscovered link. Recognizing that we can redescribe the defining equation of the information content using the threshold of probability enabled the discovery of that link. After finding the link, we strictly conceptualized several things. The Weber–Fechner formula is similar in structure to the information content formula, and this

similarity invites an analogous interpretation. Consequently, the definition of the information content for physiological response was chosen to be in analogy with the definition of the information content which had been given by Shannon; the unit of the response information content is bits. Based on the analogy between the defining equation of the information content and the formula describing the Weber–Fechner law, the author was able to theorize about the digitization of this law. On the one hand, the response information content has some features of conventional information theory. On the other hand, the response information content has the different features from conventional information theory. We also defined the perception entropy which means the average amount of the human response to external physical-stimuli; the unit of the perception entropy is bits/response. Perception entropy is also very information-theoretic, in the sense that its definition originates in the combination of information theory and the Weber–Fechner law. Amazingly, we could judge the superiority of sensory organs of humans by perception entropy.

Neural coding treats the microworld such as neurons, so computations in neural coding are complicated. Our theory targets the Weber–Fechner law which is a macroscopic phenomenon, so computations in our theory are simple. In order for our theory to be firmly established, we need to elucidate the relationships between microscopic and macroscopic theories on stimuli and responses. This relationships can be implemented by determining the values of the *K*s and elementary durations. If our theory turns out to be of great usefulness, a new field of research will be created.

Then, that new field will find applications in many diverse areas such as physiology and medicine. It will be good for us to name this new field 'perceptional information theory' or 'perceptual information theory'.

## **References**

- Berg RE, Stork DG (2005) The physics of sound. Pearson Education, San Francisco
- Borst A, Theunissen FE (1999) Information theory and neural coding. Nat Neurosci 2:947–957. https://doi.org/10.1038/14731
- Burke-Gaffney MW (1963) Pogson's scale and Fechner's law. J Roy Astron Soc Can 57:3–8
- Dimitrov AG, Miller JP (2001) Neural coding and decoding: communication channels and quantization. Network: Comput Neural Syst 12:441–472. https://doi.org/10.1080/net.12.4.441.472
- 5. Fechner GT (1860) Elemente der psychophysik. Breitkopf und Härtel, Leipzig
- Halliday D, Resnick R, Walker J (2014) Waves—II. In: Principles of physics,
  10th edn. John Wiley & Sons, Singapore, pp 423–453
- Hearnshaw JB (1992) Origins of the stellar magnitude scale. Sky Telesc 84:494–499

- 8. Hopkin B (1996) Tuning systems & pitch layouts. In: Musical instrument design. See Sharp Press, Tucson, pp 21–28
- Houstoun RA, Shearer JF (1930) Fechner's law. Nature 125:891–892.
  https://doi.org/10.1038/125891b0
- Ikeda S, Manton JH (2009) Capacity of a single spiking neuron channel.
  Neural Comput 21:1714–1748. https://doi.org/10.1162/neco.2009.05-08-792
- 11. Kinsler LE, Frey AR, Coppens AB, Sanders JV (2000) The acoustic wave equation and simple solutions. In: Fundamentals of acoustics, 4th edn. John Wiley & Sons, New York, pp 113–148
- Koechlin E, Summerfield C (2007) An information theoretical approach to prefrontal executive function. Trends Cogn Sci 11:229–235.
   https://doi.org/10.1016/j.tics.2007.04.005
- MacKay DM, McCulloch WS (1952) The limiting information capacity of a neuronal link. Bull Math Biophys 14:127–135.
   https://doi.org/10.1007/BF02477711
- Mayer A (1986) On the history of the stellar magnitude scale. J Am Assoc
  Var Star Obs 15:283–285
- 15. Nutting PG (1908) The derivation of Fechner's law. Science 27:510–511. https://doi.org/10.1126/science.27.691.510
- 16. Pogson NR (1856) Magnitudes of thirty-six of the minor planets for the first day of each month of the year 1857. Mon Not Roy Astron Soc 17:12–15. https://doi.org/10.1093/mnras/17.1.12

- 17. Rapoport A, Horvath WJ (1960) The theoretical channel capacity of a single neuron as determined by various coding systems. Inf Control 3:335–350. https://doi.org/10.1016/S0019-9958(60)90917-7
- Schulman E, Cox CV (1997) Misconceptions about astronomical magnitudes.
  Am J Phys 65:1003–1007. https://doi.org/10.1119/1.18714
- Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech
  J 27:379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x
- 20. Stein RB (1967) The information capacity of nerve cells using a frequency code. Biophys J 7:797–826. https://doi.org/10.1016/S0006-3495(67)86623-2
- 21. Ward J (1876) An attempt to interpret Fechner's law. Mind 1:452–466. https://doi.org/10.1093/mind/os-1.4.452
- 22. Weber EH (1851) Der tastsinn und das gemeingefühl
- 23. Werner G, Mountcastle VB (1965) Neural activity in mechanoreceptive cutaneous afferents: Stimulus-response relations, weber functions, and information transmission. J Neurophysiol 28:359–397. https://doi.org/10.1152/jn.1965.28.2.359
- 24. Young AT (1990) How we perceive star brightnesses. Sky Telesc 79:311–313