Pain Psychophysics: 
Its Role in Measuring, Validating, and Understanding Pain

Daniel Algom

RESUMO

Psicofísica da dor: O seu papel na medida, validação e compreensão da dor
Defende-se que os fenómenos designados pela expressão “o puzzle da dor” constituem exceções a uma correspondência regulada entre estímulos e respostas. O trabalho pioneiro de Hardy, Wolff e Goodell é passado em revista, salientando-se a forma como antecipou desenvolvimentos tanto ao nível da medida liminar e supra-liminar da dor, como dos esforços para dissociar a sensação de dor da reacção de dor. A investigação actual da dor beneficia além do mais da aplicação de ferramentas psicofísicas avançadas, tais como a teoria da detecção de sinal, os modelos de medida funcional e a psicofísica da memória. Estes métodos e rotinas analíticas abrem novos caminhos à compreensão das múltiplas facetas da dor.

PALAVRAS-CHAVE: Escala do; Limiar; Medida funcional; Puzzle da dor; Psicofísica da memória; Validação de escala.

Attempts at the quantification of pain face special difficulties. The difficulties arise from the obvious association of pain sensations with a host of emotional, motivational, and cultural factors. The influence of such factors is undeniable. Their complex interactions form a mainstay of the current literature on pain. However, an unfortunate consequence of the focus on the clinical and personality facets of pain is a tendency for the unique, bizarre, and paradoxical phenomena to dominate reviews of the field, the pertinent discussions often subsumed under the term “the puzzle of pain” (Melzack, 1973). A second consequence of the disproportional attention to the “puzzle of pain” is the paucity of discussion devoted to the measurement of the magnitude of perceived pain. Such measurement has been considered to be complex, difficult, if not downright impossible. Yet another collateral drawback has been the limited role assigned to psychophysics in assaying pain. Prominent experts have challenged the very foundations of pain psychophysics, dubbing it “farical” in its “wild variability,” and denying a relation
between stimulus intensity and pain intensity (Wall, 1979). A pall hangs over pain psychophysics and, indeed, over the entire edifice of pain measurement.

A major goal of this article is to redress the imbalance. Surely, pain presents many conceptual and practical mysteries (Weisenberg, 1977, 1984), converging on the “puzzle of pain.” A stimulus causing intolerable pain on one occasion might go hardly noticed on another. The meaning for the person of the pain situation, the introduction of dedicated cognitive/coping processes, hypnotic induction, or the administration of placebos, all contribute to the large intra- and inter-individual differences observed. Pain can be crippling following noxious stimulation, but it may fail to occur even after extensive tissue injury. And, pain may persist after all the injured tissues have been healed or removed. To complicate further the study of pain, analgesic effects seem to differ in clinical and laboratory settings: morphine is more effective in reducing clinical pain than laboratory pain, but the opposite holds with respect to placebos. This much and more is well known and well documented. However, the difficult and puzzling aspects of pain should not serve to obscure the lawful dependence of the magnitude of pain sensation on the magnitude of impinging stimulus. Pain psychophysics is concerned with the discovery, elucidation, and validation of this fundamental relationship.

In order to assess the psychophysical relationship for pain, the proper stimuli and responses must be first delineated. The adequate stimulus for pain is the damaging of the tissue, often termed “noxious stimulation.” Slight and completely reversible tissue damage has often been effected in the laboratory (through heat radiation, mechanical pressure, or very loud sounds). As often applied are disagreeable, potentially noxious stimuli that come short of actual tissue damage (cold water, electric shock, or bright flashes). At the other end is the pain response, the conscious report of the individual about the severity of the experienced pain. This point needs to be elucidated given the current preoccupation with the physiological and neural substrates of pain and with the pertinent biological reactions associated with pain. It cannot be overemphasized that pain is a perceptual experience or sensation, not a muscular, glandular, or neural reaction. One can record or notice a muscle twitch, blinking, withdrawal, or a range of autonomic responses, yet the verbal report is the sole genuine expression of human pain.

William Livingston, a surgeon and pioneer of the scientific study of pain was keen on stressing the perceptual nature of pain. Thus, “what counts most with my patient and what counts with me as his physician is the amount of pain he feels. When a patient needs a surgical operation and asks me to perform it, he does not ask how deeply the knife will cut, nor would he be concerned if I were to tell him that pain signals would continue to traverse his nervous system after he had gone to sleep. What he asks is, ‘How much it will hurt me?’ He is really asking how much of the inevitable tissue injury he will consciously experience as pain” (Livingston, 1953, pp. 68-9). Verbal reports by experienced observers express the consciously felt pain. Verbal reports as the end point are routinely employed in the study of the common senses and they are similarly used in pain to yield reproducible data.

Equipped with the definition of the adequate stimulus and resulting response, there is considerable evidence showing that “wild variability” marks the data of the common senses more than those of pain. Representative reviews (Rollman, 1992; Weisenberg, 1977) reveal an impressive uniformity in pain sensitivity and perception comparable to those observed in other sensory modalities. In fact, the lawful dependence on the
stimulus of pain sensation often surpasses that measured in the other senses (Rollman & Harris, 1987). In audition, thresholds for pure tones had a standard deviation of 5.7 to 10.7 dB for frequencies between 80 Hz and 15 KHz (Dadson & King, 1952). For vibration, sensitivity on the finger tip had standard deviations of at around 10 dB (Goff, Rosner, Detre & Kennard, 1965). In vision, threshold for a group of dark-adapted participants varied over a 4-fold range (Hecht & Mandelbaum, 1948). And, olfactory sensitivity spanned a 20-fold range even after correcting the data for stimulus noise and errors of measurement (Rabin & Cain, 1986).

In a dedicated study, Ippolitov (1972) found a 28-fold range in scotopic visual threshold with the most consistent persons from a large group of observers. He found a 15.5 dB range for pure tone threshold, a 3.5-fold range in mechanical pressure threshold using von Frey hairs, and a 6-fold range in electrotactile threshold for constant voltage stimulation. Rollman and Harris (1987), who reviewed the pertinent studies in audition, vision, olfaction, and somesthesia, themselves found a 8-fold range for pain threshold induced by constant current stimulation, and a similar range for pain tolerance (a value at which no stronger stimuli are acceptable to the participant). Clearly, pain data tend to exhibit regularity to a greater extent than do those collected by students of the common senses. Below, I show that the first rigorous measurements of pain thresholds in man (Hardy, Wolff & Goodell, 1943) yielded even more remarkably uniform results.

Let us take a final look at the puzzle of pain while considering the regular stimulus-response mapping revealed in pain psychophysics. The disproportional focus on the puzzle has clouded view of the fundamental psychophysical relationship underlying virtually any pain experience. Livingston (who incidentally was Melzack’s most influential teacher, cf. Warga, 1987) concluded that “In the majority of instances pain is proportional to the injury” (Livingston, 1953, p. 64, emphasis supplied). In everyday life, pain intensity is proportional to the force of the blow, the heat of the iron, the pressure of the sound, or the depth of the wound. This ubiquitous mapping is the reason why people are surprised when their pain differs noticeably from what they would have expected. Thus, “We wonder why some insignificant-looking scar should give severe pain, or why serious injury is not noticed in the excitement of an automobile accident” (Livingston, 1953, p. 64; see also Livingston, 1943). It is the lawfulness of the stimulus-response relationship that enables the puzzle of pain to emerge in the first place. In the absence of a proportional pain sensation there would not have been a puzzle to ponder. The exceptions subsumed under the puzzle are important, but one should recall that they are exceptions nonetheless.

As alluded earlier, the first serious attempt at the psychophysical measurement of pain was a most meritorious one, a project unsurpassed for its rigor, elegance, and ingenuity. The project spanned a period of 15 years, roughly between 1940 and 1955, and was conducted jointly by Thomas Hardy, Harold Wolff, and Helen Goodell. The authors typically served as their own subjects, sustaining months-long periods of daily radiations of heat, pressure to the skin, shivering in cold booths and sweating in intolerably hot rooms, undergoing sleep deprivation through 24-hour periods of enforced wakefulness, or enduring stress, frustration, and anxiety through (threats of) unpredictable multiple pains among a range of other stressors. Through their dedication and talent they succeeded in establishing the foundations of the edifice of pain psychophysics.
The Pain Psychophysics of Hardy, Wolff, and Goodell

The pain-generating apparatus developed by Hardy, Wolff, and Goodell was the Dolorimeter. The device focused the heat of a 500 – or a 1000 – watt lamp on a small area of the skin, usually that of the forehead or the volar surface of the forearm. The intensity of the radiation was controlled first by a rheostat and later by a transformer along with a sensitive voltmeter especially adapted to indicate intensity in milli calories per second per square centimeter. The projector also had a shutter to control time of exposure. Typical exposure duration per stimulus was 3 seconds.

Hardy, Wolff and Goodell (1943) applied this form of noxious stimulation in their first attempt to measure the pain threshold in man. They used a version of the classical psychophysical method of limits; on each successive presentation, the threshold was approached by changing the stimulus intensity by a small amount until the boundary sensation was reached. The threshold thus measured was determined in a group of 150 individuals. The group included men and women from 14 to 74 years of age. The pain threshold was found remarkably uniform. The average threshold was 220 milli calories per second per square centimeter with an inter-individual variation of a mere 5%. This small figure applied intra-individually, too. Applying successive observations on a single individual, her or his threshold could be reproduced within 5% of the previous (or of the cumulative) value. These values proved exceptionally consistent in subsequent measurements on additional samples (Hardy, Wolff, & Goodell, 1952a; Livingston, 1953). It is therefore safe to say that the average normal person begins to feel pain when the heat applied to the skin reaches 220 mcal/sec/cm², and that for most people the threshold is within 5% of that figure.

Hardy, Wolff, and Goodell were not naive to think that sensation is the sole determinant of the pain experience. They were certainly not blind to the influence of emotional and motivational factors. Consequently, these authors have carefully distinguished between pain sensation and pain reaction and suggested avenues for dissociating the two. Pain sensation refers to the intensity of the pain experience measurable by tools developed within pain psychophysics. Pain reaction refers to emotional and physiological changes in response to the pain sensation, a reaction “which apparently varies widely from person to person and in the same person under different conditions” and includes “feelings identified by the subjects as fear, anger, anxiety and many subconscious bodily changes contained in the emergency response” (Furer & Hardy, 1950, pp. 72-73). Most important therefore is the “separation of the pain experience into two components, the sensation proper and the reactions to the sensation” (Furer & Hardy, 1950, p. 72). Careful delineation of the two aspects of pain forms a major theme of the classic volume by Hardy, Wolff and Goodell (1952a), Pain sensations and reactions, the classification (and title!) conveniently overlooked by many subsequent investigators.

In order to effect the dissociation, already in their 1943 study Hardy, Wolff and Goodell posed the question. Is there a dependence of the pain threshold upon the emotional state of the individual? To answer this question, the pain threshold was determined daily over a period of several months in three individuals (i.e., the authors). On each day prior to the measurements, the individual completed a statement regarding her or his general effectiveness and mood. Although during this period the estimates of mood and efficacy showed wide variation, the pain thresholds did not, remaining remarkably constant from day-to-day.
Does threshold perception of pain depend on fatigue? To answer the question, the authors remained sleep deprived for periods of 24-hours during which pain thresholds were measured at 2-hour intervals. Although subjective well-being, fatigue, and alertness naturally varied during the times of enforced wakefulness, the pain thresholds, individually and across participants, remained stable.

For another estimate of pain reaction and its influence on pain perception, the authors recorded the change in the electrical resistance of the skin when (another portion of) the skin was exposed to heat radiation. Skin resistance is a classic measure of alarm, anxiety, and stress. Although the pain threshold was constant over a period of two months, the “alarm” reaction varied widely from participant to participant and in the same participant from time to time. The authors concluded that this outcome “demonstrates again the necessity of distinguishing between perception of pain and a reaction often associated with a disturbing experience” (Hardy, Wolff, & Goodell, 1943, p. 6).

These null effects should not be taken to mean that the pain threshold is completely immune to effects of variables other than intensity. Distraction and suggestion can alter (elevate) the pain threshold. Hardy, Wolff, and Goodell (1943) tested the effect of distraction separately by (a) reading an adventure story, (b) repeating five to nine digits forward and backward, or (c) clanging a loud bell, concurrently with the experienced pain. They further tested the effects of “autosuggestion” (having the person convince herself that she could not feel pain) and hypnosis. The authors found that each of these manipulations acted to raise the pain threshold (approximately by 15% on average).

In summary, the pain threshold in man is relatively uniform and stable. It is independent of age, sex, emotional state, and fatigue within the limits, of course, of maintaining proper attentional functioning. Elevation of the pain threshold can occur due to psychological factors of distraction and/or suggestion, due to damage to (portions of) the nervous system (a missing nerve cannot convey pain signals) or due to the action of analgesic agents. These effects granted, under ordinary conditions “the threshold for the perception of pain is of the same order of uniformity as is the pulse rate or the number of white blood cells” (Hardy, Wolff, & Goodell, 1943, p. 5).

The threshold studies and results by Hardy, Wolff and Goodell have survived generations of subsequent scrutiny and remain a pillar of modern pain psychophysics. However, the valuable work on threshold does not nearly exhaust the contributions of Hardy, Wolff, and Goodell to pain psychophysics. They extended their purview into studying suprathreshold levels of pain, developing the first scientific scale of pain, the dolor scale. I now outline this important development.

The authors first set out to determine the effective stimulus range for (heat induced) pain. It is well known that the parallel range for brightness of flashes or for loudness of tones is over 10 billion fold. For pain, the stimulus range is extremely limited, approximately 2- 3-fold at most. For their heat stimuli, Hardy, Wolff, and Goodell (1947) measured a maximum value of approximately 500 mcal/sec/cm². The limited range is dictated by the nature of the noxious stimulus impinging on the particular sensory surface.

The threshold value of 220 mcal/sec/cm² entails an amount of heat that will redden the skin after repeated tests, and it is fairly close to the level of heat at which cells are irreversibly damaged (cf. Livingston, 1953). Increasing the intensity of the radiation augments the magnitude of pain of the same pricking or burning nature. Once a fairly
intense value has been passed, a stimulus that burns the skin, the pain actually lessens or changes because the burning process destroys sensory fibers in the skin and deeper tissues have fewer such fibers. Discriminability decreases at this point due probably to the altered nature of the sensation, too. The sensation at the burning level (or at a value just preceding the burning level) can be called “ceiling” pain. Indeed, Hardy et al (1947) reported that “The sensation evoked by a stimulus of about 680 mcal/sec/cm² is therefore a ‘ceiling’ pain since stimuli of greater intensity cause no perceptibly greater pain” (p. 1154), and that discrimination beyond that level is virtually impossible.

Determining thus the effective stimulus range, Hardy et al (1947) proceeded by parsing this range onto equally discriminable steps. Applying again a version of the method of limits, the authors determined all values of the difference threshold or the difference limen, DL, from absolute threshold to ceiling. According to Weber’s law, the stimulus ratios marking successive DLs, $I_1/I_0$, $I_2/I_1$, ... $I_n/I_{n-1}$, are equal. Each DL marks a just noticeable difference (JND) in sensation. Hardy et al (1947) found 21 JNDs for detecting minimal increments in the magnitude of pain sensation from threshold to ceiling. Fechner (1860) assumed that every (equally noticed) difference between stimuli separated by a JND is equal in sensation (note that Weber merely stated the equality of the successive stimulus ratios, and did not make assumptions about the associated sensations or sensation increments). Subsequently, using Weber’s law in tandem with his own principle that equally noticed differences are equal in sensation, Fechner derived his famous psychophysical law. The law states that sensation magnitude is proportional to the logarithm of stimulus intensity (see Marks & Algom, 1998, and Marks & Gescheider 2002, for extensive discussions). Hardy et al (1947) espoused Fechner’s law to derive their dol scale of pain.

![Figure 1 – The dol scale of pain. One dol is equal to two successive JNDs. (After Gescheider, 1997).](image-url)
Figure 1 depicts the dol scale. It is based on the cumulative number of JNDs from threshold as a function of stimulus intensity. The dol, a Fechnerian unit of measurement of pain, is equal to 2 JNDs. The dol scale is a good demonstration how valid psychophysical scales can be constructed by obtaining DLs experimentally (the Weber fraction is not constant over the entire stimulus range), then adding up the DLs along the stimulus continuum.

Of course, a scale is as good as are its assumptions. Hardy, Wolff, and Goodell (1948) proceeded to test those assumptions, anticipating methods of modern psychophysics. As I just recounted, the dol scale is based on Fechnerian principles and derivations. The question can be posed, Is the scale consistent with ratio properties? Does the sum of 2 JNDs (1 dol) represent the same difference in pain intensity in all parts of the dol scale? Is pain in dols associative? Is a 4-dol pain twice as intense as a 2-dol pain? Hardy, Wolff and Goodell (1948) tested these questions using the method of "fractionation," a forerunner of the method of magnitude estimation (Stevens, 1975). At the beginning of the experiment, the participant was exposed to a standard stimulus. Then, the participant was asked to report the intensity of pains evoked by subsequent stimuli in terms of fractions of the initial, standard pain. The scale erected on the basis of the fractionation data coincided remarkably well with the dol scale (obtained by summation of successive DLs). In earlier research by Wolff, patients used a method in which pain intensity was indicated by the number of "pluses" entered on an arbitrary scale (with one representing threshold and ten the most intense pain ever experienced). Wolff's method is actually that of magnitude estimation with a bounded range. The results obtained by Wolff's method also agreed well with the dol scale.

The authors concluded that the numbers on the dol scale, representing the intensity of pain, "are capable of being added and divided according to the ordinary rules of arithmetic and it is possible to define one dol both as the sum of two JNDs and as one-tenth of the intensity of the ceiling pain. It is proper, also, to speak of one pain as being either 'twice as strong' as another or as equivalent in intensity to the sum of two smaller pains" (Hardy, Wolff, & Goodell, 1948, p. 385). In modern psychophysics, summated JND scales often fail tests of consistency (Zwislocki & Jordan, 1986), although sometimes they pass them (Heinemann, 1961; see Marks & Algom, 1998, for further discussion). Concerning the dol scale in particular, over much of the range it was found to be proportional to a scale derived on the basis of prototypical magnitude estimation (Adair, J. C. Stevens, & Marks, 1968).

In order to test the generality of the dol scale, Hardy, Wolff and Goodell (1952b) derived the scale on the basis of another type of experimental pain, an aching pain induced by applying pressure to the skin. The effective range of this form of noxious stimulation lies between 400 and 6000 grams, although the authors did not use values beyond 2000 gram in the experiments. The first DL above threshold averaged at around 125 grams. Clearly, this stimulus differs noticeably from the original heat stimulus in terms of both range and DLs. In order to test the validity of the dol scale, the authors used the method of cross modality matching (CMM) in the first known application of that procedure. To avoid anachronism, CMM was introduced and named by Stevens (1975), but Hardy et al anticipated the method by years. The participant first received a stimulus of the original thermal radiation type causing pricking pain. Then, the participant received several stimuli of the pressure type causing aching pain. From the latter, the
participant "selected those which in his opinion were equal in intensity to the pricking pain previously experienced... The reverse of the above procedure was accomplished by first presenting an aching pain as standard to be matched in intensity by a series of pricking pains" (Hardy, Wolff, & Goodell, 1952b, p. 251). The CMM function derived by Hardy et al (1952b) is presented in Figure 2.

![Figure 2](image.png)

Figure 2 – Cross-modality matching between the perceived intensity of aching pain produced by pressure and the perceived intensity of pricking pain produced by heat radiation. The diagonal represents pains of equal numbers of dols on the respective continua. (After Hardy et al. 1952).

Note that the data in Figure 2 are plotted in terms of the JNDs above threshold on the respective pain dimensions. The units differ as a matter of course, mcal/sec/cm² for thermal radiation, grams for pressure. The pain threshold for pressure is approximately 500 gram, whereas the threshold for heat is approximately 220 mcal/sec/cm². The respective DLs above threshold are 125 gram and 14 mcal/sec/cm², respectively. Nevertheless, when the pains are measured in terms of the respective numbers of JNDs above threshold or dols, then comparisons become possible and meaningful. It is then seen that pricking pain and aching pain are judged to be of equal intensity when they are evoked by stimuli of the same number of JNDs above threshold. Therefore, the dol scale can be used extensively to measure virtually any form of noxious stimulation, experimental or natural.

Finally, Furer and Hardy (1950) tested the role of emotional and motivational factors in suprathreshold pain (heat radiation). During a six-months period, the participants were exposed to painful radiation on a daily basis while measuring their galvanic skin response. To avoid habituation of that response or to reinstate it, the participants were exposed to various forms of stress and were also studied in cold (7 deg C) and hot (49 deg C) environments. In contrast to the uniform accuracy with which pains were perceived, there was marked variation in the galvanic skin response. Despite the variation, the magnitude of the galvanic skin response increased in proportion to the intensity of the experimental pain. This relationship might contribute to the failure by
many to properly distinguish pain sensation from pain reaction. The galvanic skin response is a measure of the threat content of a pain (the greater the pain the greater the threat) and thus only indirectly related to pain intensity. Notably, threat can be posed by a variety of means other than intensity. Notably, too, repeated presentation of a (strong) stimulus reduces the galvanic skin response although the pain remains virtually unchanged.

It is difficult to fully appreciate the fundamental contributions of Hardy, Wolff, and Goodell to the psychophysics of pain within the confines of a brief review. They first measured pain thresholds in man with unprecedented sophistication and accuracy. They first applied the basic discoveries of Weber and Fechner to erect a premier scientific scale of pain. And, they exhibited conceptual clarity by carefully dissociated pain sensations and reactions. To a brief look at subsequent attempts at accomplishing this dissociation I turn next.

Dissociating Perceptual and Cognitive Components of Pain by Methods of Sensory Decision Theory

To a first impression, the methods of sensory decision theory or the theory of signal detectability (TSD) seem singularly well suited for separating sensory and response bias components in pain (Green & Swets, 1966; Macmillan & Creelman, 2004). The first component entails purely perceptual processes associated with pain intensity. The second component encompasses the gamut of emotional, motivational, and cognitive factors that are also associated with the pain experience. Because the latter components play such a prominent role in pain, their separation from the former is highly desirable. TSD provides a means of independently measuring the two components in an observer’s detection performance.

However, applying the TSD routine to pain presents special difficulties. TSD experiments in the common senses pursue detection processes, focusing on the confusion between signal and noise. Pain, in contrast, requires intense levels of stimulation; blank trials are impractical. To create the uncertainty needed to tap the various measures (TSD requires less than perfect discriminability), pain researchers have presented pairs of intense stimuli. Observers were asked to discriminate between those stimuli. In other cases, observers were asked to rate several level of noxious stimulation, the ratings serving to derive receiver operation characteristic curves which in turn served to determine sensitivity and response bias.

Therefore, in a prototypical experiment the observers rate the painfulness of several levels of noxious stimulation before, during, and after the administration of an analgesic agent such as acupuncture. Applying this procedure, Clark and Yang (1974) found that observers assigned lower numbers to painful thermal stimuli after acupuncture. Moreover, they found that the sensitivity parameter (for discriminating adjacent levels) did not change after acupuncture, but that the criterion parameter was higher. The authors concluded that acupuncture had no sensory effects; it merely causes subjects to raise their criteria when reporting pain. Consequently, Clark and Yang (1974) attributed the analgesic properties of acupuncture to cognitive changes. Similar reasoning (with varying outcomes) is found throughout the literature (Rollman, 1992).
Rollman (1977; see also Rollman, 1992) reviewed portions of the pertinent literature and found it marked with inconsistencies, methodological inadequacies, and questionable interpretation of the data. Of more consequence, in a brilliant analysis Rollman showed that TSD is inherently incapable to accomplish the sought separation in pain. Thus, "signal detection theory methods fail to permit an unequivocal separation between changes in pain responsiveness caused by sensory modulation and those due to changes in criterion or response bias. Both can appear to produce alterations in the 'criterion' parameter. Moreover, even when the 'sensitivity' parameter is reduced, it is impossible to conclude that a true analgesic effect has occurred, since all one can say with certitude is that the stimuli are less easily distinguished. They may be less painful, equally painful, or even more painful than before treatment" (Rollman, 1992, p. 539).

In order to understand Rollman's analysis, let us examine the two situations depicted in Figure 3, those before and after modulation by an analgesic agent. Clearly, the administration of the analgesic agent reduced neural activity; the distributions following modulation have been shifted to the left along the sensation continuum. Note, too, that the analgesic agent reduced activity produced by both the weaker and stronger stimulus. This uniform reduction left the distance between the stimuli, and the sensitivity parameter, unchanged. Traditional interpretation by TSD leads to erroneous conclusions in this situation. Because sensitivity did not change, an unseasoned observer might conclude that the sensory effects remained similarly unchanged. However, this is patently not true because the neural activity has been modulated (i.e., pain was reduced) for both stimuli following modulation. The sensitivity parameter reflects the ability to discriminate between the stimuli, not their level of painfulness.

![Figure 3 - Reduction of neural activity in both the weaker and the stronger of two noxious stimuli after the administration of an analgesic agent. The pain is reduced following modulation but the discrimination index remains unchanged. The criterion parameter is increased although the criterion remained constant. (After Rollman, 1977).](image-url)

Note also that, following modulation, the criterion parameter is increased even though the criterion has not changed! The distributions of the pair of noxious stimuli have shifted to the left with the criterion remaining at a constant value. This leads to the
spurious outcome of a seeming criterion shift. The upshot is clear: TSD is inherently incapable of accomplishing the necessary separation between pain sensation and pain reaction.

Reviewing the TSD-induced "revolution" in pain research, one must respect the inchoate ideas and attempts by Hardy, Wolff, and Goodell (1952a,b). The time-honored adage that experimental resolution is preferable to conceptual or statistical one is reaffirmed.

The reader might have noticed that all the studies discussed up to this point, early and recent, entailed a single continuum of stimulation (be it radiant heat, pressure, electrical current, or cold water) various levels of which were presented for judgments of painfulness. Given the one-dimensional designs, the question of validity remains ultimately unresolved. Although control of the experimental stimuli has been sufficiently tight (physics and engineering provide increasingly better methods for measuring environmental energies), it cannot be known with certitude that the person's observable reports accurately reflect the pains felt. Clinical pain, arising spontaneously in natural settings, seems to enjoy better ecological validity. After all, most normal persons seek health care, certainly pain reduction or relief, when they actually suffer from pain (Weisenberg, 1977; see also Beecher, 1956). However, clinical studies are naturally lacking at the stimulus end; the irritating stimulus (headache, low-back pain, toothache, or an aching stomach) cannot be objectively measured. Available information is typically confined to the complaint of the individual. Can experimental and clinical pain nonetheless be combined by way of validating human pain reports?

Psychophysics in the Field: Measuring the Pain of Labor Contractions

Under ordinary circumstances, laboratory pain is relatively free of influence by emotional factors. Although this feature can be utilized to probe aspects of pain sensation, it is actually of dubious value. It hampers generalization and reduces ecological validity. The capacity of experimental pain to control for the spatiotemporal aspects of the stimulus is hardly advantageous if results cannot be generated beyond the laboratory. Generalization is problematic precisely because clinical pain is fraught with emotion. Clinical pain produced by pathological processes is inevitably linked to the individual's concerns about her or his well-being. The laboratory presents a different context, lacking the anxiety associated with disease, the threat of disfigurement, or death (Weisenberg, 1977). The labor of childbirth offers a unique opportunity to combine the advantageous features of laboratory and clinical pain by way of enhancing validity.

Pain of labor arises naturally from an endogenous source and it is laden with the full spectrum of stressful emotions, hope, and anxiety. It is thus a foremost example of clinical pain. However, the labor of childbirth also carries a precious feature of experimental pain, namely, the possibility of objective measurement of the irritating stimulus. Although the irritating stimuli are entirely natural and spontaneous (i.e., not under experimental control), they are nonetheless measurable via standard biometric equipment. Labor pain thus combines advantageous features of both laboratory pain (measurability) and clinical pain (natural source and spontaneous arousal), offering a unique avenue for valid assessment.
Algom and Lubel (1994) took advantage of this situation, measuring pain during labor. They had women estimate the painfulness of labor contractions. Each woman judged the intensity of pain felt at each of five contractions selected by the experimenter (naturally, a different set of contraction pressures with each woman). The intensities of the uterine contractions selected and judged were biometrically measured (intruterine pressure in mmHg). The magnitude estimates of pain of the women were then plotted as a function of peak pressure of their uterine contractions. The individual pain scales agreed fairly well with power functions. The average exponent was 1.39. It follows that for one contraction to feel twice as painful as another its intensity would need to be merely 1.64 times as great. Clearly, the pain of labor contractions grows at a faster rate than does the intensity of the pressure stimulus that generates the pain. Why does (labor) pain grow as a positively accelerated function of the stimulus? The rapid growth rate probably serves to bring on the experience of intense pain rapidly as stimulus intensity increases and approaches the point of harm or injury to the person. This arrangement “is adaptive because the rapid welling up of intense pain causes the individual to take action to reduce pain and avoid serious injury” (Gescheider, 1977, p. 321).

The Algom and Lubel study enjoys ecological validity. It elucidated the characteristics of a self-generated, endogenous pain in its natural habitat. These unique conditions conferred validation support. Nevertheless, even this study cannot rule out completely nonlinear influences on the over judgments. One-dimensional designs are structurally unsuitable to resolve the issue of validity. Multidimensional designs are better equipped to tackle questions of validity.

**Multidimensional Designs: Functional Measurement of Pain**

Norman Anderson’s functional measurement approach (Anderson, 1981, 1982, 1992) shows the value of studying the algebraic structures that underlie information integration in sensory processing. The approach prescribes the use of (at least) two physical continua, whose levels vary in a factorial design. The data generated by the multidimensional design permit to derive the rule that governs the integration of information from the two stimulus continua. Often, integration is found to obey simple algebraic rules such as addition or averaging. Notably, this rule is predicated on or supported by substantive theory that exists in the field of interest. In this way, substantive theory precedes measurement and confers validation support on measurement.

If the overt responses (ratings, magnitude estimates) reproduce the structure and values predicted by the theory-induced integration rule, than those data can be taken as joint evidence for both the integration rule and the validity of the response scale. Once the responses are validated, one can derive reliable psychophysical functions for each of the stimulus dimensional.

Algom, Raphaeli, and Cohen-Raz (1986, 1987) applied Anderson’s logic and methods to the study of pain. The authors covaried the values of two separate noxious variables in a factorial design. They combined 6 levels of electric current applied to the wrist with 6 levels of uncomfortably loud tones fed in the ears, making 36 tone-shock compounds in all. Subjects gave magnitude estimates of the painfulness of these
concurrently-presented stimuli. Figure 4 reproduces the results of the 1986 study. Each curve represents a different fixed intensity tone. Clearly, the louder the tone, the greater the pain. As clearly, pain similarly increased with increasing shock current.

Figure 4 – Additive integration of pain. Mean magnitude estimates of pain plotted as a function of current intensity delivered to the wrist. Each curve represents a constant SPL delivered to the two ears. (After Algom et al. 1986).

The most salient characteristic of the factorial plot is the roughly equal spacing of the family of curves in the vertical dimension (although a slight trend toward divergence in the upper right is evident). Parallel spacing implies linear additivity of the numerical responses. Therefore, the aversiveness of an electric shock and a loud tone presented simultaneously at various intensities approximates the linear sum of the individual painful components. Note, too, that the shock only trials (bottom curve) have the same slope as the as do the compounds of shock plus tone (the same was true of the tone-only trials). This features supports and additive composition rule over an averaging rule.

Observed parallelism is consistent with three features of the functional measurement model. First, the integration of information follows the simple algebraic rule of addition (supporting for this set of data Algom's functional theory of pain; Algom, 1992a). Second, the judgment functions appear linear: The overt judgments mirror the underlying additive rule. Third, by implication, the psychophysical functions are valid. Accordingly, Algom et al. derived psychophysical functions separately for shock-induced pain and for acoustically induced pain. In fact, a pair of functions was derived for each dimension: one was based on single-component presentations (shock- or tone-only trials), the other on the marginal means calculated across the values of the other dimension for each level of the relevant dimension. The psychophysical functions approximated power functions with exponents of about 1.1 for shock and 0.9 for sound. Note that the function for auditory pain differs appreciably from the functions for loudness that are routinely derived for acoustic sounds.
In a study conducted by T. Meidler in the author's laboratory, testing was extended to ternary compounds. Levels of three noxious variables were varied in a full $5 \times 5 \times 5$ factorial design. Subjects made magnitude estimations of the painfulness of compounds made up of an electric shock, a loud tone, and a bright flash (to dark adapted eyes). The results are presented in Figure 5. Salient to visual inspection is the roughly parallel spacing of the surfaces. The absence of a three-way interaction in the appropriate ANOVA supports the additive structure. The estimates of overall pain approximated the sum of the individual electric, acoustic, and visual painful components. Concerning the respective psychophysical functions, pain related differently to the three inducing stimuli. Notably, the pain functions for tones and flashes differed noticeably from the well-known sone and brill scales that relate loudness to sound pressure and brightness to luminance.

![Figure 5](image)

**Figure 5** - Additive integration of electrical, acoustical, and visual pain. Average magnitude estimates are plotted as a function of SPL delivered to the two ears and flashes of light presented to the two (dark adapted) eyes. Each surface represent a constant current delivered concurrently to the wrist.

The application of multivariate designs represents clear progress in studying and validating pain. Anderson's approach is particularly promising, harnessing substantive theory in the quest to validate scales of pain. However, there is an indeterminacy in Anderson's model, too. A trade-off between the pattern of integration and the judgment function permits different models to account for the same set of data (with appropriately chosen judgment functions). Anderson devised methods (e.g., two-operation designs) to enable one to reject some alternatives, but thus far no general solution has been provided (Gescheider, 1977; Gigerenzer & Murray, 1987; Marks & Algorn, 1998; Marks &
Gescheider, 2002). If there is a way to verify the assumed integration rule in an independent fashion, then the method is fully valid and offers an even more valuable tool for studying pain and other senses.

Obviously, pain psychophysics focuses on the perception of concurrent pain. So do the vast majority of published pain studies. However, the scope of pain is wider, and it certainly includes memory. In the clinic, often the memory of the referent pain only is available. Who cannot recall a “toothache that seemed intolerable during the night... but which had almost disappeared the next morning when [one] reluctantly climbed into the dentist’s chair” (Livingston, 1953, p. 64). Even when the pain persists, memory for pain provides valuable medical information (“When did you first notice it?” “Did you feel pain yesterday?” “If so, how intense was it?”). An extension of psychophysics, memory psychophysics or mnemophysics (Algom, 1992b; Algom & Marks, 1989) applies the arsenal of psychophysical methods to study remembered sensations. Although young, the memory psychophysics of pain has already yielded valuable data.

Memory Psychophysics for Pain

In the prototypical experiment, observers are presented with stimuli to which they learn to associate names (or colors or any other type of label). At some time later, the observer is presented with the name and required to judge the magnitude of the referent stimulus from memory. Standard magnitude estimations can be used for judgments of remembered magnitudes. In the aforementioned study by Algom and Lubel (1994), additional groups of women estimated the pain of labor contractions from memory. Each woman was required to associate colors (presented on large cardboards) with each of five uterine contractions during labor. After either 8, 24, or 48 hours from the time of labor, the woman was again presented with the colors and asked to magnitude estimate the pain she remembered to be associated with each color. The memory pain functions also approximated power functions with average exponents of 1.63, 2.0, and 1.75, respectively for pain remembered 8, 24, and 48 hours after labor. Notable is the lawful dependence of remembered pain on the intensity of the irritating stimulus. Remembered pains connect to the original stimulus through the same relation, power transform, as do perceived pains. Notable, too, are the different exponents. Those that govern remembered pain are larger than that governing perceived pain (1.39). The average perceptual and memory-based pain functions are depicted in Figure 6.

Why does remembered pain grow at an even faster rate than does perceived pain? The very expansive psychophysical function for remembered pain may comprise an advanced warning device in which remembered pain triggers withdrawal and other actions aimed at avoiding further injury to a part of the body that has already sustained damage and pain.

Pain psychophysics can also serve to resolve theoretical questions of general interest. It has been generally found that the exponents for remembered stimuli are smaller than those for perceived stimuli (Algom, 1992b). For example, perceived area relates to physical area by a power function with an exponent of 0.64, whereas remembered area relates to physical area by a power function with an exponent of 0.46. Two hypotheses have been proposed to account for this finding. According to the
reperception hypothesis, stimulus intensity is first transformed into sensation magnitude according to Stevens' power law. Thus,

$$S = aI^b$$  \hspace{1cm} (1)

where $S$ is sensation magnitude and $I$ is stimulus intensity. Of the constants, $a$ is a measurement constant and $b$ is the exponent. For memory, the same transformation is applied again but this time it acts on the perceptually transformed values $S$ rather than on the original physical values $I$. Thus,

$$M = a'S^b$$  \hspace{1cm} (2)

where $M$ is the respective memorial magnitude, and $a'$ has the same meaning as $a$. Substituting for $S$ via equation 1 yields

$$M = AI^{bd}$$  \hspace{1cm} (3)

where $A$ is the new scaling factor. Therefore, the reperception hypothesis predicts that the exponent for memory should be the square of the exponent for perception. Indeed, for area, the memory exponent (0.46) is close to the square of the perception exponent (0.64).

According to the alternative uncertainty hypothesis, observers experience greater uncertainty in making judgments from memory than in making perceptual judgment. They thus restrict the range of their estimates (Kerst & Howard, 1978) or widen the effective stimulus range (Algom, Wolf, & Bergman, 1985), resulting in the reduction of the exponent.

How does one decide between these alternatives? Both theories predict smaller memory exponents for perceptual dimensions governed by smaller-than-unity exponents, the generally observed phenomenon. Different predictions exist only for dimensions that are characterized by larger-than-unity exponents in perception. For such expansive
dimensions, the predictions differ. The reperception hypothesis predicts a larger memory exponent, whereas the uncertainty hypothesis predicts a smaller memory exponent regardless of what happens in perception. The Algom and Lubel pain study comprised the critical experiment needed for theoretical resolution. Consistent with the reperception hypothesis and inconsistent with the uncertainty hypothesis, the exponents of the memory functions were found larger than those of the perceptual functions.

Memory psychophysics becomes perhaps most powerful when applied with multidimensional designs. Then, rules of integration in memory can be examined in tandem with the psychophysical function and compared with the parallel values in perception. Functional measurement in particular can be profitably applied to study organizational principles in memory for pain. Z. Lindenberg, in the author’s laboratory, had separate groups of observers judge the painfulness of compounds of electric shocks and disagreeably loud tones (see, Algom et al., 1986). One group of participants judged painfulness upon administration of the shock-tone mixtures, replicating the perceptual conditions of Algom et al. (1986, 1987). The participants in a second group judged the painfulness of the same stimulus mixtures from memory. The participants in the third group judged mental shock-tone mixtures. The subject first learned names to separately presented shocks and tones (i.e., shock-tone compounds were not presented). Later, presented with a pair of names, the subject imagined the appropriate noxious compound (that he or she has never experienced) and gave an estimate of its painfulness. The results are presented in Figure 7.

![Figure 7](image)

**Figure 7** - Integration of pain in perceptual, remembered, and mental compounds of painful electrical shocks and disagreeably loud tones. The perceptual and mental judgments exhibit additivity although remembered judgments not.

The additive structure evident in the perceptual estimates replicates that found by Algom et al. (1986, 1987). Additivity failed, however, for the remembered compounds. The judgments from memory obeyed a nonlinear rule of concatenation. Why do people add perceived pains, but multiply the same pains when they refer to them from memory? Can affective factors alter remembered pain? The multiplicative structure acts to magnify certain pains (particularly the strongest compounds) perhaps as a device of advanced warning against repeating a very painful experience. Finally, the additive structure reappeared in mental compounds of pain. The integration rule for imagined pain – events constructed wholly subjectively – was linear summation of the individual pains. People
act on imagery just the way they do facing real pain. People thus carry a core of ecologically valid knowledge about pain that is probably cognitively impenetrable. They know more about pain that they can consciously articulate. People's world of make-believe, sustaining imagery wounds or other mishaps, is governed by the same combinational principles that form their physically-bound sensory world.

RESUMÉ

Psychophysique de la douleur: Son rôle dans la mesure, validation et compréhension de la douleur

On montre que les phénomènes regroupés sous l'expression “puzzle de la douleur” sont des exceptions à la correspondance régée entre stimulus et réponses. On passe en revue le travail innovateur de Hardy, Wolff et Goodell, en soulignant la façon dont il devance des essors postérieurs aussi bien dans la mesure liminaire et supraliminaire de la douleur que dans les efforts pour départir entre sensation de douleur et réaction de douleur. La recherche actuelle sur la douleur bénéficie en plus de l'emploi d'outils psychophysiques avancés, tels que la théorie de la détection de signal, les modèles de mesure fonctionnelle et la psychophysique de la mémoire. Ces méthodes et routines analytiques ouvrent des avenues nouvelles à la recherche d'une compréhension plus poussée des multiples facettes de la douleur.

Mots-clés: Échelle dol; Seuil; Mesure fonctionnelle; Puzzle de la douleur; Psychophysique de la mémoire; Validation d'échelle.

ABSTRACT

Pain Psychophysics: Its Role in Measuring, Validating, and Understanding Pain

The paradoxical phenomena subsumed under the “puzzle of pain” are shown to be exceptions to a consistent stimulus-response mapping. The pioneering work of Hardy, Wolff, and Goodell is reviewed, anticipating modern developments in threshold and suprathreshold measurement as well as in the attempts to dissociate pain sensation and reaction. Modern pain research further benefits form the application of advanced tools of psychophysics including signal detection theory, models of functional measurement, and memory psychophysics. These methods and analytical routines open up new avenues in the quest to understand the multiple facets of pain.

Key-words: Dol scale; Threshold; Functional measurement; Puzzle of pain; Memory psychophysics; Scale validation.

BIBLIOGRAFIA


Furer, M., & Hardy, J. D. (1950). The reaction to pain as determined by the galvanic skin response. *Life Stress and Bodily Disease*, 29, 72-89.


Algom, D.


