tified, these differences can lead to the selection of anesthetic-analgesic regimes that are safe and effective in human infants.

Our special issue closes with *A Question of Pain in Invertebrates* by Dr. Jane Smith. Dr. Smith reviews the physiologic and behavioral responses seen in certain invertebrates submitted to noxious stimuli. She briefly discusses the organization of the nervous systems of several invertebrate families and presents evidence for both nociceptors and a system of endogenous and behavioral responses to repeated noxious stimuli. Despite these observations, it is still difficult to demonstrate conclusively that cephalopods actually perceive pain. This being the case, Dr. Smith concludes that invertebrates should be given "the benefit of the doubt where questions of pain and suffering are concerned." She provides some practical guidelines for investigators dealing with invertebrates.

During the past several years, a variety of meetings and publications have been devoted to the issue of assessment and alleviation of pain in experimental animals. In an upcoming issue of *ILAR News* we will present summaries of: the ILAR/NRC report entitled *Recognition and Alleviation of Pain in Laboratory Animals*, the UFAW document *Alleviation of Pain in Laboratory Animals*, the Cornell University International Symposium on *Animal Pain and Its Control* and the 1990 ACLAM forum dealing with *Pain and Distress in Animals*. Additional articles relevant to the issue of animal pain and distress will be published in subsequent issues of *ILAR News*. Your comments on this issue would be appreciated by the editorial panel.

---

**Pain in Animals and Humans**

**Behavioral Assessment of Pain: Nonverbal Measures in Animals and Humans**

Francis J. Keefe, Ph.D., Roger B. Fillingim, Ph.D., and David A. Williams, Ph.D.

**INTRODUCTION**

Pain has been defined as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage" (International Association for the Study of Pain Subcommittee on Taxonomy, 1979). Given that pain is a subjective experience, it is perhaps not surprising that so much effort in human pain assessment is devoted to the measurement and analysis of pain perception. Pain, however, can have very significant effects on the behavior of humans and animals. In clinical settings, patients experiencing pain often display a variety of behaviors that communicate the fact that they are experiencing pain (Fordyce, 1976). These pain behaviors may include complaints of pain, reductions in activity, increased medication intake, or alterations in facial expressions or body posture. Animals who are exposed to painful stimuli are likewise known to exhibit pain behaviors such as withdrawal from the painful stimulus or other self-protective maneuvers.

Since pain behaviors are overt, they can be observed and recorded. Systematic assessments of pain behaviors have long been viewed as useful in analyzing pain phenomena in animals and are increasingly viewed as valuable by human pain researchers (Dubner, 1989). In the past 2 decades, standardized methods for carrying out behavioral assessments in humans and animals have been developed and evaluated by a number of researchers. Research carried out by these investigators has demonstrated that objective and reliable measures of behavior can be quite useful in the analysis of pain phenomena.

The purpose of this paper is to provide the reader with an overview of behavioral assessment methods commonly used in evaluating pain in animals and humans. The paper reviews behavioral assessment methods utilized in three major settings: (1) the animal laboratory, (2) the...
human laboratory, and (3) clinical research and practice. For each setting, we describe common pain induction and pain measurement methods, review the strengths and weaknesses of these methods, and highlight important practical and ethical issues.

**ANIMAL LABORATORY MEASURES**

**Pain Induction Methods**

The goal of many nonhuman laboratory pain tests is to measure the presence or absence of analgesia following some experimental manipulation (e.g., surgery, drug administration, or exposure to stress). If the methods of pain induction are equivalent, then any change in pain can be attributed to the experimental manipulation rather than to the pain induction procedure. Pain induction procedures allow experimenters varying degrees of control over the parameters of the pain stimulus (e.g., intensity, duration, location, and temporal patterning). Prominent pain induction methodologies include the tail-flick test, the hot-plate test, the pinch test, and the formalin test. Other pain induction procedures, such as the adjuvant-induced model of arthritis, seek to mimic human forms of clinical pain in hopes of gaining a better understanding of the pathophysiologic basis of disease-related pain phenomena.

The tail-flick test was first described by D’Amour and Smith (1941). When radiant heat is directed to the tail of a rat, a stable, spinally mediated nociceptive response occurs in the form of a flicking or jerking of the tail away from the heat source (see Williams and Thorn, 1984 for details on apparatus and procedure). Baseline latencies to remove the tail from the heat are established by adjusting the heat intensity to the point where the average latency to flick the tail is constant for four consecutive trials. Thereafter, the heat source remains constant for the remainder of that testing session. New baselines are established with each new testing session. Differences between the experimental and baseline latencies are interpreted as an index of analgesia. Increases in the latency for the rat to flick its tail are indicative of analgesia, while decreases in tail-flick latency are indicative of hyperalgesia.

The expression of the tail-flick reflex has been shown to be modulated by the administration of neurohormonal agents or opioid drugs, or through activation of descending pathways from the brain to the spinal cord (Akil and Mayer, 1972; Akil and Liebeskind, 1975; Yaksh and Wilson, 1979). The sensitivity of this reflex to chemical modulation makes this test particularly attractive for analgesic substance testing.

The hot-plate test was first described by Woolfe and Smith (1941). When radiant heat is directed to the tail of a rat, a stable, spinally mediated nociceptive response occurs in the form of a flicking or jerking of the tail away from the heat source (see Williams and Thorn, 1984 for details on apparatus and procedure). Baseline latencies to remove the tail from the heat are established by adjusting the heat intensity to the point where the average latency to flick the tail is constant for four consecutive trials. Thereafter, the heat source remains constant for the remainder of that testing session. New baselines are established with each new testing session. Differences between the experimental and baseline latencies are interpreted as an index of analgesia. Increases in the latency for the rat to flick its tail are indicative of analgesia, while decreases in tail-flick latency are indicative of hyperalgesia.

The hot-plate test was first described by Woolfe and Smith (1941). When radiant heat is directed to the tail of a rat, a stable, spinally mediated nociceptive response occurs in the form of a flicking or jerking of the tail away from the heat source (see Williams and Thorn, 1984 for details on apparatus and procedure). Baseline latencies to remove the tail from the heat are established by adjusting the heat intensity to the point where the average latency to flick the tail is constant for four consecutive trials. Thereafter, the heat source remains constant for the remainder of that testing session. New baselines are established with each new testing session. Differences between the experimental and baseline latencies are interpreted as an index of analgesia. Increases in the latency for the rat to flick its tail are indicative of analgesia, while decreases in tail-flick latency are indicative of hyperalgesia.

The expression of the tail-flick reflex has been shown to be modulated by the administration of neurohormonal agents or opioid drugs, or through activation of descending pathways from the brain to the spinal cord (Akil and Mayer, 1972; Akil and Liebeskind, 1975; Yaksh and Wilson, 1979). The sensitivity of this reflex to chemical modulation makes this test particularly attractive for analgesic substance testing.

The hot-plate test was first described by Woolfe and Smith (1941). When radiant heat is directed to the tail of a rat, a stable, spinally mediated nociceptive response occurs in the form of a flicking or jerking of the tail away from the heat source (see Williams and Thorn, 1984 for details on apparatus and procedure). Baseline latencies to remove the tail from the heat are established by adjusting the heat intensity to the point where the average latency to flick the tail is constant for four consecutive trials. Thereafter, the heat source remains constant for the remainder of that testing session. New baselines are established with each new testing session. Differences between the experimental and baseline latencies are interpreted as an index of analgesia. Increases in the latency for the rat to flick its tail are indicative of analgesia, while decreases in tail-flick latency are indicative of hyperalgesia.

The expression of the tail-flick reflex has been shown to be modulated by the administration of neurohormonal agents or opioid drugs, or through activation of descending pathways from the brain to the spinal cord (Akil and Mayer, 1972; Akil and Liebeskind, 1975; Yaksh and Wilson, 1979). The sensitivity of this reflex to chemical modulation makes this test particularly attractive for analgesic substance testing.

The hot-plate test was first described by Woolfe and Smith (1941). When radiant heat is directed to the tail of a rat, a stable, spinally mediated nociceptive response occurs in the form of a flicking or jerking of the tail away from the heat source (see Williams and Thorn, 1984 for details on apparatus and procedure). Baseline latencies to remove the tail from the heat are established by adjusting the heat intensity to the point where the average latency to flick the tail is constant for four consecutive trials. Thereafter, the heat source remains constant for the remainder of that testing session. New baselines are established with each new testing session. Differences between the experimental and baseline latencies are interpreted as an index of analgesia. Increases in the latency for the rat to flick its tail are indicative of analgesia, while decreases in tail-flick latency are indicative of hyperalgesia.

The expression of the tail-flick reflex has been shown to be modulated by the administration of neurohormonal agents or opioid drugs, or through activation of descending pathways from the brain to the spinal cord (Akil and Mayer, 1972; Akil and Liebeskind, 1975; Yaksh and Wilson, 1979). The sensitivity of this reflex to chemical modulation makes this test particularly attractive for analgesic substance testing.

The hot-plate test was first described by Woolfe and Smith (1941). When radiant heat is directed to the tail of a rat, a stable, spinally mediated nociceptive response occurs in the form of a flicking or jerking of the tail away from the heat source (see Williams and Thorn, 1984 for details on apparatus and procedure). Baseline latencies to remove the tail from the heat are established by adjusting the heat intensity to the point where the average latency to flick the tail is constant for four consecutive trials. Thereafter, the heat source remains constant for the remainder of that testing session. New baselines are established with each new testing session. Differences between the experimental and baseline latencies are interpreted as an index of analgesia. Increases in the latency for the rat to flick its tail are indicative of analgesia, while decreases in tail-flick latency are indicative of hyperalgesia.

The expression of the tail-flick reflex has been shown to be modulated by the administration of neurohormonal agents or opioid drugs, or through activation of descending pathways from the brain to the spinal cord (Akil and Mayer, 1972; Akil and Liebeskind, 1975; Yaksh and Wilson, 1979). The sensitivity of this reflex to chemical modulation makes this test particularly attractive for analgesic substance testing.

The hot-plate test was first described by Woolfe and Smith (1941). When radiant heat is directed to the tail of a rat, a stable, spinally mediated nociceptive response occurs in the form of a flicking or jerking of the tail away from the heat source (see Williams and Thorn, 1984 for details on apparatus and procedure). Baseline latencies to remove the tail from the heat are established by adjusting the heat intensity to the point where the average latency to flick the tail is constant for four consecutive trials. Thereafter, the heat source remains constant for the remainder of that testing session. New baselines are established with each new testing session. Differences between the experimental and baseline latencies are interpreted as an index of analgesia. Increases in the latency for the rat to flick its tail are indicative of analgesia, while decreases in tail-flick latency are indicative of hyperalgesia.

The expression of the tail-flick reflex has been shown to be modulated by the administration of neurohormonal agents or opioid drugs, or through activation of descending pathways from the brain to the spinal cord (Akil and Mayer, 1972; Akil and Liebeskind, 1975; Yaksh and Wilson, 1979). The sensitivity of this reflex to chemical modulation makes this test particularly attractive for analgesic substance testing.

The hot-plate test was first described by Woolfe and Smith (1941). When radiant heat is directed to the tail of a rat, a stable, spinally mediated nociceptive response occurs in the form of a flicking or jerking of the tail away from the heat source (see Williams and Thorn, 1984 for details on apparatus and procedure). Baseline latencies to remove the tail from the heat are established by adjusting the heat intensity to the point where the average latency to flick the tail is constant for four consecutive trials. Thereafter, the heat source remains constant for the remainder of that testing session. New baselines are established with each new testing session. Differences between the experimental and baseline latencies are interpreted as an index of analgesia. Increases in the latency for the rat to flick its tail are indicative of analgesia, while decreases in tail-flick latency are indicative of hyperalgesia.

The expression of the tail-flick reflex has been shown to be modulated by the administration of neurohormonal agents or opioid drugs, or through activation of descending pathways from the brain to the spinal cord (Akil and Mayer, 1972; Akil and Liebeskind, 1975; Yaksh and Wilson, 1979). The sensitivity of this reflex to chemical modulation makes this test particularly attractive for analgesic substance testing.
ness is used during observation to prevent them from lying down.

Various paradigms have been developed to mimic chronic pain conditions in humans. These paradigms typically involve the creation of mock chronic pain conditions by surgical manipulation or by injection with or exposure to various diseases. For example, the adjuvant-induced arthritis model is a method of producing polyarthritis by intradermal injection of killed *Mycobacterium butyricum* (Newbould, 1963; Gouret et al., 1976; De Castro Costa et al., 1981). Although various rat strains may react differently to the effects of the adjuvant, a preclinical phase of 15 days is generally sufficient before clinical signs of arthritis appear.

Other examples of chronic pain models include trigeminal neuralgia (Black, 1974), neumora (Chuder and Dong, 1983), central pain and itch syndromes (Kryzanovsky, 1976), autotomy (Duckrow and Taub, 1977), visceral pain (Ness and Gebhart, 1990), phantom limb pain (Levitt and Heybach, 1931) and general chronic pain models (Lombard et al., 1979; Albe-Fessard and Lombard, 1983).

**Pain Measurement Methods**

The behavioral expression of nociception varies depending upon the evolutionary advancement of the animal under study, the type of pain stimulus used, and any endogenous or exogenous factors that might modulate the pain response (Kavaliers, 1988). In order to quantify pain in nonhumans, numerous independent measures of pain behavior have been developed. With the proliferation of pain behavior measures comes the problem of adequate reliability and validity of these pain measures. Frequently observational systems are developed for a single study and the psychometric properties of the system go unassessed. Even with commonly used procedures, slight differences exist in how the behavior is assessed from study to study. The lack of standardization introduces error into the results of these studies and makes interpretation and comparison across studies difficult. This section will review several of the more commonly used and researched methods of nonhuman pain behavior observation.

Analgesia tests, such as the tail-flick test, seek to observe a single, predictable reflex. The presence of the reflex can be recorded by human observers or even by mechanical or photoelectric cells. Other means of measuring pain assess more complex constellations of pain behaviors and require more skillful human observers and elaborate methods of measurement.

To quantify the presence of analgesia, formulas such as the Percent Maximum Potential Analgesia (% MPA) have been developed (Watkins and Mayer, 1982). To quantify tail-flick analgesia, % MPA is expressed by the formula, % MPA=\((TL-BL)/CT-BL\)\) x 100, where TL is the test trial tail-flick latency, BL is the baseline tail-flick latency, and CT is the cut-off time. Cut-off time is typically considered to be two times the baseline latency and is used to prevent tissue damage to an analgesic animal. Percent MPA can be similarly calculated for other analgesia tests such as the pinch test by substituting latency with measures of force (see Hayes et al., 1978).

The formalin test (Dubuisson and Dennis, 1977) has associated with it one of the more elaborate methods of observing and rating nonhuman pain. There are two separate observation systems (one for rats and one for cats) that are scored on a zero- to three-point rating scale. Higher ratings are associated with greater displays of pain behavior (e.g., licking, biting, shaking the injected paw). This behavioral rating system shows evidence of good reliability and validity.

Behavioral observation systems of pain in animal models of human pain conditions need to be sensitive to behaviors that are specific to the particular pain conditions under study. For example, Butler et al. (1985) provided a quantifiable evaluation scale for assessing the clinical signs and pain behaviors of arthritis in rats. Clinical signs included: gross swelling and deformity of the hind paws, forepaws, and toes; deformity, swelling, and ulceration of the tail; inflammation of the penis; conjunctivitis; and nose bleeding. Pain behaviors include scratching, limited mobility and exploring, and vocalization. Scratching was defined as a complex motor response beginning with licking of a body part, then changing to biting of the skin, pulling of the fingers or toes through clenched teeth, or vibrating the hind paw near the ear. Exploring encompassed rearing, sniffing, running, and climbing. A mobility score (scaled 0=rat lies down only, to 4=rat walks and runs normally) provided an assessment of functional ability. These authors also employed a Randall-Selitto apparatus (a method of applying a graded weight on the hind paw) so as to determine the threshold for vocalization quantified in units of force.

**Comment**

This section on animal laboratory measures describes some of the more commonly used methods of experimental pain induction and measurement. Although 10 years ago many of the methods described might have been used interchangeably, we now know that each test has advantages and limitations that are influenced by evolutionary development, the various anatomical pain systems, and suggestibility to confounding endogenous factors.

Many of the analgesia tests (e.g., tail-flick, pinch test) are confounded by the need to restrain the animal during testing. Biphasic changes in nociception, both antinociceptive and hyperalgesic, have been associated with restraint (Pilcher and Browne, 1983; Jorum, 1988; Porro and Carli, 1988). Pain induction methods, such as the flinch-jump or formalin test, that permit unrestrained motor move-
ment are less affected by this limitation. One problem inherent in each of these pain induction methods is that they involve handling of the animal. Simple handling of rats has been shown to recruit endogenous opioid responses that can influence response latencies (Jorum and Shyu, 1988).

Associated with the limitation of restraint, are limitations associated with the duration of the testing period. One advantage of the tail-flick test is the relative speed and ease by which this test can be completed. Tests such as the hot-plate test or flinch-jump appear to be measuring highly integrated escape behaviors suggestive of higher cortical involvement. Complex motor behaviors can be considerably more variable in their expression than spinal reflexes thus requiring considerably more trials in order to establish a stable baseline. Mayer and Liebeskind (1974) reported sessions lasting as long as 2 hours per animal for the flinch-jump test.

Recent studies have shown that several pain systems influence pain perception; however, these systems are not equally distributed throughout the body or equally responsive to the same temporal patterns of stimulation. For example, opioid systems appear to have a greater influence on nociception for forepaws when delivered in brief bouts. In contrast, hind paws appear to be more nonopioid mediated (Watkins et al., 1982). Methodologies that assess only one body region, such as the tail-flick test, might not be sensitive to the complexities of nociception throughout the body. The pinch test, which assesses nociception from five different body regions, might better address this limitation.

The duration of the nociceptive stimuli also influences the behavioral expression of pain and should be considered when evaluating pain behavior. Intermittent, long duration stressors appear to be more opiate mediated than short-term stressors (Amit and Galina, 1986, 1988). Thus the tail-flick test, being a short-term stressor would be better suited for testing phasic models of pain than for testing chronic models of pain. Alternatively, the formalin test offers an analogue of a more enduring pain that might better mimic naturally occurring pain. In addition this methodology has the advantage of allowing the experimenter to observe the animal over a period of an hour (e.g., walking, sitting, standing).

Animal models of human pain conditions permit one to study the development of pain behaviors as a disease process progresses. Such models are especially good for studying and testing various treatments that can eliminate or retard the onset of pain-related behaviors. A limitation of all the methods mentioned is attendant to the suffering that the test animal experiences as a result of experimental manipulation or pain-induction procedure. As in any experimental procedure involving pain, the decision to induce pain must be carefully weighed against the advances in knowledge that can be gained. This is particularly important in the case of nonhuman pain research where informed consent is not possible. It is the ethical responsibility of all researchers using animals in the search for knowledge to minimize the amount of suffering animals experience (Loeb et al., 1989).

HUMAN LABORATORY MEASURES

Pain Induction Methods

Laboratory research for the study of pain in humans has historically utilized numerous pain induction methods. Beecher (1959) identified several requirements for an ideal painful stimulus, including: (1) it must cause minimal tissue damage and hazard to subjects, (2) there must be a relationship between the stimulus intensity and the degree of pain experienced, (3) it must be repeatable without altering responses over time, (4) it must be sensitive, (5) it must be easily applied and produce distinct painful sensation, (6) it must show analgesic dose relation, (7) it must be applicable to both humans and animals, and (8) it must produce natural sensations. Subsequent criteria (Gracely, 1984, 1985) required that the pain induction have a rapid onset and termination, similar sensitivities across individuals, and a stimulus that excites primarily nociceptive afferents. Major pain induction methods in humans include mechanical pressure, chemical stimulation, ischemic pain, cold pressor pain, electrical stimulation, and thermal stimulation (Gracely, 1989).

In the past, mechanical pressure has been applied at various anatomic locations, including the finger, mastoid process, esophagus, and bile duct (Beecher, 1959). Recent studies have generally used pressure applied to a finger joint (Whipple and Komisurak, 1985). This stimulus produces natural pain with a wide range of intensities and durations; however, the rapidity of pain onset and termination is not easily controlled. Also, this method is likely to activate a wide range of afferent neurons that convey extraneuronal information.

Chemical pain stimulation can be achieved by application of an irritant to an exposed anatomic surface (e.g., blistered skin or nasal mucosa) or by intramuscular injection (Beecher, 1959; Kobal, 1985; Foster and Weston, 1986). Like mechanical stimulation, chemically induced pain can be natural and intense, but stimulus control is suboptimal.

The ischemic pain test is conducted by inflating a pressure cuff on the upper arm of a subject who then opens and closes his/her hand at a fixed rate (Smith et al., 1966). Although adequate reliability and validity of this technique have been reported, it has been found to be insensitive to analgesic manipulation (Sternbach, 1983). This widely used stimulus produces a natural, continuous, and increasing pain, but onset and termination are gradual, and there appears to be a practice effect with repetition (Sternbach, 1983).

Cold pressor pain is produced by the immersion of a
Pain Measurement Methods

Dubner (1985) and Gracely (1983) have identified important properties of pain measurement methods for use in laboratory studies of human pain phenomena. They suggest that any method used should be (1) sensitive to changes in stimulus intensity, (2) sensitive to changes throughout the range from threshold to tolerance, (3) sensitive to manipulations that alter sensory discriminative capacities of the subject, (4) free from subject and experimenter bias, (5) able to separate intensity from quality of pain, and (6) characterized by absolute rather than relative scales to allow use across both time and individuals. Several nonverbal techniques are frequently used to measure responses to laboratory pain stimuli in humans. These include visual analog scales (VAS), electromyography (EMG) and other psychophysiological measures, electroencephalography (EEG), and facial expressions.

A visual analog scale is a line, usually 10 cm long that represents a continuum of pain experience. The endpoints generally represent the limits of the pain experience (e.g., no pain vs. severe pain). The subject is requested to mark the line at a point corresponding to the severity of the pain being experienced. With the exception of the explanation, the VAS is independent of language, has been found to be quite reliable, and seems more sensitive than simple descriptive scales (Huskisson, 1983). Like all self-reporting scales, VAS are susceptible to subject and experimenter bias. Also, inadequate explanation or lack of understanding by the subject can lead to error.

Psychophysiological measures, such as EMG activity, cardiovascular reactivity, and skin conductance, are often used in the laboratory assessment of human pain. Findings are difficult to summarize because there is wide variability across subjects, and responses in many of these physiologic systems might be attributable to the emotional changes, such as anxiety, that often accompany pain. Nonetheless, some general conclusions can be tentatively drawn. In response to pain, gastrointestinal motility decreases; respiration increases; muscular activity in the region of stimulation increases; systolic and diastolic blood pressure typically increase; and augmented heart rate and stroke volume as well as peripheral vasoconstriction are sometimes observed (Sternbach, 1968). The major drawbacks of these measures are their nonspecific relation to pain and relatively rapid habituation over time (Bromm and Scharein, 1982).

Recently, investigators have examined the utility of cerebral evoked potentials elicited by painful stimuli as a pain assessment method. Evoked potentials are highly correlated with subjective pain reports; they are sensitive to the effects of analgesic drugs, and appear to differentiate first pain (A-delta fiber mediated) from second pain (C-fiber mediated) (Bromm, 1989). Additionally, this measure has low susceptibility to experimenter and subject bias. However, there is habituation in the response over time and pain can be reported in the absence of a definable evoked potential response (Bromm and Scharein, 1982; Harkins, et al., 1983).

Facial expressions in response to pain can be measured using the Facial Action Coding System (FACS) (Ekman and Friesen, 1978). The FACS is a reliable method of scoring the frequency, duration, and intensity of facial expressions from videotaped recordings. In a recent study of laboratory pain, the most consistent facial response to pain was blinking or closing the eyes (Craig and Patrick, 1985). These authors reported extreme interindividual variability, and their results were somewhat inconsistent with previous descriptions of prototypical facial expressions of pain (Hjortsjö, 1969). While facial expressions may be less susceptible to subject bias
(Chapman and Jones, 1944), it is unclear which expressions are specific to pain, since facial expressions vary as a function of emotional state.

One interesting, but relatively infrequently used measure of pain response in humans is the nociceptive reflex measure. This technique involves recording electromyographic activity in targeted muscles following electrical stimulation. The method is reliable across time and individuals, and the magnitude of EMG responses has been correlated with both the intensity of electrical stimulation and self-reported pain. Nociceptive reflex measures have also been found to be sensitive to the effects of analgesic medication (DeBroucker et al., 1989). However, the nociceptive reflex and subjective pain experience can be dissociated; therefore, this measure should not be considered an objective measure of pain (Price, 1989).

Comment

Selection of both pain induction and pain measurement methods for human laboratory research should be based on practical and ethical concerns as well as conceptual and theoretical considerations. Important practical issues include expense, equipment and personnel requirements, and feasibility for use in a particular experimental paradigm. Ethically, the pain induction method should involve the least possible discomfort and risk of injury, and the use of a painful stimulus must be scientifically justifiable. Additionally, subjects must be thoroughly informed of the expected level of discomfort and the potential risks and benefits of their participation in the research. Conceptual and theoretical considerations include not only the experimental questions to be addressed, but also previous empirical or clinical research relevant to these questions. Researchers interested in extrapolating their findings to clinical settings may want to adopt a pain induction technique that produces sensations mimicking clinical pain (e.g., ischemic pain, mechanical pain). Also, employing pain measurement methods analogous to those used clinically would be ideal. On the other hand, researchers wishing to examine the use of various manipulations on repeated pain experiences over a short time period must use pain stimuli with quick onset and offset and requiring minimal practice.

Although the above discussion indicates that choice of pain induction and measurement methods will depend on multiple factors, a few general guidelines can be suggested. First, if examining a particular treatment or manipulation, use pain induction and measurement techniques that have been employed with similar treatments and manipulations in the past. This permits comparisons of new results with previous findings. Secondly, it is preferable to use multiple pain assessment techniques than to use only one. For example, a VAS and an EEG used to examine pain responses provide information from separate systems that are differentially susceptible to experimenter and subject bias. Additionally, different measures can be validated against each other in the same experiment. Finally, when choosing pain induction strategies, the requirements mentioned above must be considered. In this vein, all things being equal, laser delivered radiant heat is probably the most desirable pain induction method. Its intensity can be quantified, it produces natural pain, it produces measurable first and second pain, and it appears to be highly comparable across individuals. Its major flaw is its relatively slow termination.

CLINICAL MEASURES

Pain Induction Methods

Clinical pain assessment is a much more complex and difficult task than laboratory pain assessment. When dealing with clinical pain, one usually cannot quantify the tissue pathology basis of pain or directly measure the duration or intensity of the pain related to this tissue damage. In the laboratory, in contrast, the experimenter not only controls the magnitude and duration of pain stimuli but can also control for environmental factors such as temperature, sound level, and light that might affect pain behavior.

In order to better understand behavior patterns in patients experiencing clinical pain, researchers have taken two major approaches (Keefe, 1989). The first approach involves measuring patient behavior in naturalistic settings. Repeated measurements of salient behaviors are taken over long time periods and variations in behavior are studied. A patient having low back pain, for example, might be asked to keep a daily diary record of time spent up and out of the reclining position for 4 weeks prior to starting treatment. When naturalistic recording is used there is no attempt to induce clinical pain. Instead, one examines the frequency or duration of pain-related behaviors and attempts to identify factors that may explain variations in behavior.

The second behavioral approach to assessing pain is to sample behavior in structured situations that are likely to elicit pain. A patient having neck pain, for example, may be escorted to an examination room and asked to engage in a series of timed neck flexion, extension, and rotational movements and observed for evidence of guarded movement or pain-related facial expressions. Such structured situations can be conceptualized as a pain induction method. Although they do not provide the precise control over noxious stimulation and environmental factors that can be attained in the laboratory setting, they do have the advantage of introducing some standardization in terms of behavioral demands and environment into the process of clinical pain assessment.
Pain Measurement Methods

Nonverbal measures used in behavioral assessment can be grouped into three broad categories: diary methods, electromechanical devices, and observational methods.

Diary methods. Behavioral assessment approaches to studying clinical pain have long relied on diary methods that are designed to measure such important pain-related behaviors as activity level and medication intake. A variety of diary formats have been developed and many of these are modeled after the activity diary record developed by Fordyce (1976). This record consists of a sheet on which patients are asked to make hourly entries noting (1) the amount of time they spend sitting, reclining, standing or walking; (2) medication intake; and (3) pain intensity on a zero- to 10-point scale. The data gathered are usually graphed, and patterns of activity are identified. One common pattern, which we call the pain cycle (Gil et al., 1988), is evident when a patient persists with activities until he reaches the point of pain tolerance, and only then allows himself to rest and take pain medication. As Fordyce (1976) noted, rest and pain medications may serve to reinforce and maintain this maladaptive behavior pattern. This cycle can be broken using two techniques: (1) interspersing periods of moderate activity with limited rest periods and (2) placing medications on a time-contingent rather than pain-contingent schedule of delivery (Fordyce, 1976; Gil et al., 1988).

Follick et al., (1984) validated an activity diary method designed for use with chronic pain patients. They developed their daily diary format to record activity patterns including position, time spent alone or with others, time at home, medication intake, and use of a variety of pain relieving methods (e.g., hot pack, heating pad, or transcutaneous electrical nerve stimulator). The diary was also used to gather information on ratings of pain, mood, and tension. Patients using this diary received detailed written and verbal instructions and were asked to make entries in the diary three times daily. Data analysis revealed that the diary was quite reliable. Reliability estimates indicated that patients’ pain-related activity patterns and pain ratings were relatively consistent from day to day. In addition, high correlations were obtained between patient and spouse recordings of activity level and medication intake. Finally, patient reports of time spent up and out of the reclining position (uptime) were highly correlated (r=0.94, p<0.01) with records obtained using automated electromechanical uptime monitors worn by the patients. Taken together, these findings indicate that daily activity diaries can provide reliable and valid measures of patient behavior in the natural environment.

Blanchard and Andrasik (1985) developed a headache diary that is now widely utilized in clinical and research settings. Patients are asked to rate the level of their headache activity four times daily and to indicate the medications taken for the headache. The level of pain is rated on a six-point scale in which zero equals no headache and six represents an extremely intense, incapacitating headache. Scores derived from the diary include a headache index, or an average daily headache measure; a measure of headache free days; and a peak headache rating, which is the highest rating obtained for a given time period. Medication intake is quantified using a method developed by Coyne et al., (1976). This method scales the potency of medications enabling one to calculate a medication index by multiplying the potency scale value by the number of doses. Research suggests that the headache diary used by Blanchard and Andrasik (1985) possesses good reliability and validity and is sensitive enough to detect improvements occurring during biofeedback or relaxation treatment programs (Blanchard et al., 1981).

Electromechanical Devices. Electromechanical devices provide another means of measuring activity level in patients experiencing pain. The earliest approaches to electromechanical measurement relied on commonly available devices such as clocks or pedometers. One example is an “uptime clock” developed by Cairns et al., (1976). The clock automatically recorded the amount of time spent out of bed by patients hospitalized on a special pain management unit. A microswitch fitted to the patient’s bed triggered the clock mechanism each time the patient was out of bed. Another example is the use of pedometers to record distance walked by pain patients daily (Saunders et al., 1978).

Sanders (1980) modified the timing unit in an electronic calculator to measure directly the amount of time patients experiencing low back pain spent standing or walking. This method relies on a microswitch placed on the patient’s leg that is activated each time the patient stands or walks. Follick et al., (1985) developed and evaluated a similar uptime monitor.

Recently, there has been a great deal of interest in the actometer as an objective method of quantifying activity level (Tryon, 1985). An actometer is a self-winding watch that has been altered so that it measures movement rather than time. Basically, each time the patient moves an accelerative force is applied to the watch. This force results in movement of the second hand on the watch. One can quantify levels of movement by taking periodic measurements from the watch. We carried out a series of studies to evaluate the applicability of an actometer to assessment of chronic pain patients (Morrell and Keefe, 1988). Although the actometer provided readings that were quite internally reliable (i.e., readings from two yoked actometers covaried strongly with distance walked), actometer readings were quite variable over time. Patients walking the same distance on 3 separate days often showed substantial differences in actometer readings. It appeared that variations in movement patterns such as greater limping while walking one day versus another, may have accounted for the poor reliability across time. Thus, the utility of actometers...
for use with patients experiencing chronic pain is questionable.

Observation Methods. Because pain-related behaviors are overt, they also can be recorded directly in the natural environment by trained observers. Cinciripini and Floreen (1982) developed an observation method to record a variety of pain behaviors exhibited by patients experiencing chronic pain who were participating in a multidisciplinary pain management program. Their observation method was designed to assess the percentage of time patients spent daily engaging in nonverbal pain behavior, talking about pain, complaining about nonpain topics, talking about healthy topics, and engaging in assertive behaviors. A time sampling approach was used in which a trained staff member carefully watched each patient for 5 minutes of every half hour during the waking day. This observational method was found to be highly reliable and sensitive enough to detect changes in behavior resulting from treatment.

Several investigators have also used trained observers to measure pain behavior occurring in physical examination or treatment settings. In our research (Keefe et al., 1984, 1986), we directly observed and recorded five pain behaviors that patients with low back pain exhibit during physical examination—guarding, bracing, rubbing of the painful area, grimacing, and sighing. Considerable variability in pain behavior was noted. Multivariate regression analysis revealed that the best predictors of pain behavior during the physical examination were physical findings and depression. Patients having multiple positive findings had higher levels of pain behavior, as did patients scoring high on a self-report inventory of depression.

Another approach to assessing pain-related behaviors during physical examination is that developed by Waddell et al., (1980). Waddell and his colleagues developed a standardized physical examination procedure for patients with low back pain that included measures of nonorganic physical signs such as nonanatomic tenderness, abnormal weakness, and exaggerated verbal or nonverbal behaviors. Reliability studies demonstrated that independent clinicians could reliably score and record the nonorganic signs and that the signs had good test-retest reliability. Validity data indicated that these signs occurred significantly less often in patients having definite pathology of the spine than in patients having few organic physical findings. This method of recording patient behavior during physical examination has become widely utilized by orthopedists in the United States and Great Britain.

Elliott et al. (1987) have developed an observation scale for measuring children’s pain and distress during medical procedures. Trained observers used this scale to record the behavior of 55 pediatric cancer patients, ages three to 13, who were undergoing painful bone marrow aspirations. Behaviors recorded included information seeking, crying, screaming, requiring physical restraint, verbal resistance, seeking emotional support, verbalizations of pain, and flailing. Interrater reliability for the scale was found to be quite high (r=0.98) and scores on the scale were found to correlate highly with nurse ratings of distress, patient ratings of fear, and measures of heart rate.

As mentioned earlier, structured situations can be used to sample important pain-related behaviors in patients with clinical pain. We have developed a structured approach to assessing pain behavior in patients with low back pain that involves a 10 minute standardized video-taped session (Keefe and Block, 1982). In order to induce pain behavior, patients are asked to sit, stand, walk, and recline for 1- to 2-minute periods. The order of these activities is randomly determined, and each patient completes the entire set of movements. Each videotape is scored by trained observers who record the occurrence and nonoccurrence of five pain behaviors (guarding, bracing, rubbing of the painful area, grimacing, and sighing) using ongoing 20- and 10-second record intervals. Our research has shown that this observation method is both reliable and valid. Observers working independently and simultaneously have shown a high degree of agreement as to the behaviors observed (percentage agreement = 93-99 percent). The level of pain behavior also correlates highly with patient ratings of pain and with ratings of pain by individuals who are unaware of the observation procedure. The observation method also appears to be sensitive to the effect of treatment. Substantial reductions in pain behavior were noted in patients with low back pain who participated in a pain management program.

We have also used structured situations to sample pain in a variety of other chronic pain populations, including patients with osteoarthritis (Keefe et al., 1987), rheumatoid arthritis (Anderson et al., 1987), or head and neck cancer (Keefe et al., 1985). Data from these studies have provided additional support for the reliability, validity, and utility of such observation methods.

Comment

Pain diaries, electromechanical devices, and observation methods all have been used to assist clinicians in carrying out behavioral assessments of patients experiencing pain. Research studies evaluating these assessment procedures have been carried out mainly in patients experiencing chronic pain. With the exception of studies such as that carried out by Elliott, et al. (1987), the utility of behavioral methods for assessing acute pain phenomena has not been evaluated.

Of the assessment methods reviewed in this section, pain diaries are undoubtedly the most widely used. Diaries are practical, inexpensive, and easily adapted to meet the needs of a given patient or clinician. The reliability and validity of pain-diary data, however, is often ques-
tionable. Research studies have shown that patients with chronic pain often underestimate their activity levels (Sanders, 1983) and can be inaccurate in their reports of intake of pain medications (Ready et al., 1982). As Fordyce (1976) suggests, the reliability and validity of pain-diary data can be enhanced if one provides a careful rationale for diary recording, as well as some systematic training in methods of entering data on diary forms. It is interesting to note that the strongest support for the validity of pain-diary data has come from studies that have trained patients in use of diary methods (e.g., Blanchard et al., 1981; Follick et al., 1984).

Electromechanical measurement can provide objective assessments of activity in patients experiencing clinical pain. Many of these devices are not available commercially, and thus have not been widely utilized.

Observation methods show promise for behavioral assessment. These methods are generally considered to yield the most fundamental and objective measures of behavior (Keefe, 1989). Most of the methods that have been developed require extensive training of observers and are most likely to be applicable to research settings. A number of researchers, however, have been working on developing observation procedures, such as rating scales or checklists, that are simpler and more practical (Keefe, 1989). With further development of these alternative approaches, it is likely that behavioral observation will become more widely used clinically.

There are important ethical issues involved in the use of behavioral assessment methods in clinical pain settings. First and foremost is the issue of informed consent. Behavioral assessment methods should only be used when the patient has been provided with a clear explanation of the nature and purpose of the assessment method and has voluntarily agreed to the assessment. Informed consent not only fulfills ethical requirements, but also enhances the quality of data collected. A second major concern is confidentiality. Data gathered by means of direct observation or electromechanical devices is subject to the same restrictions on confidentiality that apply to any clinical data. This behavioral data should be handled with care to guard the identity of the patient and should not be released without the expressly written consent of the patient. Confidentiality is especially important when collecting data from patients with chronic pain who may be involved in litigation related to their pain. A third ethical concern is misuse and misinterpretation of behavioral assessment data (Keefe, 1989). Some clinicians, for example, view pain behaviors in an oversimplified fashion and consider any occurrence of guarded movement or facial grimacing to be evidence of malingering. In such instances the occurrence of pain behavior may be used as a justification for failing to carry out a full evaluation of the patient's pain or for refusing medical or surgical treatment (Keefe, 1989). It is incumbent on those who utilize behavioral assessment methods to educate clinicians on interpreting pain behavior data and to provide cautions regarding the dangers of drawing overly simplistic conclusions from behavioral data.

**CONCLUSIONS**

The studies reviewed in this paper suggest that behavioral assessment techniques can provide useful strategies for understanding pain. Behavioral methods have been used in animal laboratory studies of pain for many years, and a number of reliable and valid approaches for pain induction and measurement have been developed. In the human laboratory, several sophisticated techniques for measuring behavioral responses to controlled pain stimuli have been developed. These methods appear to provide promising tools for evaluating analgesic and other treatment approaches to pain. Behavioral assessment methods applicable to the study of clinical pain have only recently been developed. Further research is needed to demonstrate the practical utility of such measures. Additional studies are also needed to compare recently developed behavioral methods for clinical pain assessment to some of the sophisticated pain perception measures currently available (Gracely, 1989).

In the next 10 to 15 years there are likely to be a number of new advances in behavioral methods for assessing pain that will further increase our ability to understand and treat acute and chronic pain phenomena.

**REFERENCES**


Pain Research and Therapy, Volume 1, J. J. Bonica and D. Albe-Fessard, eds. New York: Raven Press.


---

The Assessment of Pain in the Burned Child and Associated Studies in the Laboratory Rat

Patricia F. Osgood, Ph.D.

INTRODUCTION

Pain alleviation in infants and children has long been minimal, if not totally ignored. In recent years there at last has been a growing realization of the need for controlling pain in the young (Anand and Hickey, 1987; McGrath and Unrah, 1987; Ross and Ross, 1988; Schecter, 1989; P.A. McGrath, 1990), but the pain of burned children has received almost no attention (Perry and Heidrich, 1982). However, the extreme pain often experienced by children with burn injuries has led us to seek better means of managing pain in this group.

Patricia F. Osgood, Ph.D., is assistant professor of anesthesiology (pharmacology) at the Massachusetts General Hospital, Harvard Medical School, Shriners Burns Institute, in Boston.