Tell-tale TINT: Does the Time to Incorporate into Nest Test Evaluate Postsurgical Pain or Welfare in Mice?

Miranda S Gallo,[1] Alicia Z Karas,[1] Kathleen Pritchett-Corning,[2] Joseph P Garner[3] Guy Mulder,[4] and Brianna N Gaskill[5,*]

Identifying early indicators of distress in mice is difficult using either periodic monitoring or current technology. Likewise, poor pain identification remains a barrier to providing appropriate pain relief in many mouse models. The Time to Incorporate to Nest Test (TINT), a binary measure of the presence or absence of nesting behavior, was developed as a species-specific method of identifying moderate to severe distress and pain in mice. The current study was designed to evaluate alterations in nesting behavior after routine surgery and to validate the TINT's ability to measure pain-related behavioral changes. CD1 mice undergoing carotid artery catheterization as part of a commercial surgical cohort were randomly assigned various nesting, surgery, and analgesia conditions. To provide context for the TINT outcomes, we measured other variables affected by pain, such as weight loss, food consumption, and scores derived from the Mouse Grimace Scale (MGS). Mice that had surgery were more likely to have a negative TINT score as compared with controls. All mice were more likely to fail the TINT after receiving postoperative buprenorphine, suggesting that buprenorphine may have contributed to the failures. The TINT, MGS live scoring, and scoring MGS images all loaded strongly on a single component in a principal component analysis, indicating strong convergent validity between these measures. These data indicate that the TINT can provide a quick, objective indicator of altered welfare in mice, with the potential for a wide range of uses.

Abbreviations: TINT, Time to Incorporate to Nest Test; MGS, Mouse Grimace Scale; NQS, Nest Quality Scoring; PCA, Principle Component Analysis; GLM, General Linear Model

Welfare standards and experimental results can be compromised by the biologic consequences of pain or distress.[29] Effective, species-specific pain assessment is therefore crucial to the implementation of appropriate animal pain management, a scientifically and ethically important refinement that should be included in most modern science.[33] Pain scales that incorporate the assessment of behavioral and physiologic parameters have been used to address this need[3,6,9,30,35,47] but these methods are limited by both their lack of specificity for pain and their subjectivity. Physiologic and biochemic indicators are similarly nonspecific, and can require additional, stressful animal handling, sample analysis, and expensive equipment.[1,3,8,46] Current means of assessing pain or distress are thus ill-suited to the typical time, budget, and equipment constraints present in large laboratory animal facilities, where the mouse census can number in the thousands. These factors contribute to an unmet need for an objective, fast, and inexpensive method of cage-side pain identification in mice.

Use of an ethologic framework of behavioral time-budgets is a practical strategy for identifying pain in mice. When mice fail to perform a highly motivated behavior, this suggests that they are budgeting their energy toward more biologically essential activities such as sleeping, eating, or drinking. The absence of normal, species-specific behaviors has also been suggested as a strong indicator of pain in animals.[2,8] A suitable candidate behavior for the identification of pain in mice must therefore be robust throughout the species and conserved in captivity, but enough of a luxury behavior (unlike eating) to be reliably disrupted by a painful stimulus. This alteration of behavior thus allows for earlier identification and treatment of pain.[38]

Nest building has been established as an important normal behavior in mice of both sexes, and across many genetic backgrounds.[14,37,50] Building a nest allows for thermoregulation and improved calorie conversion, while also decreasing the development of stereotypies.[15,17,20] Mice will work to gain access to nesting material, supporting nesting as a highly motivated behavior.[14] However, depending on the environmental conditions, this behavior may be altered before others that are more crucial to survival. Nest building occupies a flexible behavior niche; existing as a necessity under extreme conditions, but not necessarily within the standard laboratory environment,[4] positioning nest building as a useful identifier of wellbeing in laboratory mice.

The Time to Incorporate to Nest Test (TINT)[49,50] was developed as a tool to quickly assess pain by determining the presence or absence of gathering behavior for nest building. A small amount of additional nesting material is introduced to a mouse cage, and a positive test result occurs when the mouse retrieves the new nesting material and incorporates it into an existing nest within 10 min. Normal, healthy mice are consistently successful in the TINT, establishing a strong baseline for this rapid, simple screening technique.[50] Neither sophisticated nesting ability nor copious amounts of nesting material are

required for mice to be TINT-positive, and even strains that are considered poor nest-builders will successfully incorporate nesting material after being provided with pretesting experience.[50]

While a significant effort has been made to establish that the underlying behavior captured by the TINT is robust in laboratory mice, the test had not been fully validated as an indicator of pain. The original work developing TINT[50] demonstrated that mice fail in potentially painful situations but was not able to show that the deficit could be reversed with analgesia. In the absence of this validation, the TINT has been directly applied to identify pain or discomfort in scientific settings.[21,23] However, this test was not developed as a scientific measure of pain, but rather was intended to aid animal care personnel in determining whether animals were behaviorally abnormal after a procedure. Without further validation, it is unknown whether the test provides meaningful information about mice in painful situations. Therefore, we sought to establish whether the TINT can demonstrate discriminant (distinguish between painful and nonpainful animals) and convergent validity (do results agree with a 'gold standard' measurement).

We hypothesized that surgical pain would reduce the motivation for nesting behavior and that postsurgical mice that were not administered analgesia would be more likely to fail the TINT. We also expected that animals provided analgesia would be more likely to pass the TINT, and that analgesia would reverse a negative TINT outcome. Moreover, we predicted that a negative TINT would correlate with other measures of pain in mice such as high MGS, reduced nest scores, loss of body weight, and decreased food consumption.

Materials and Methods

The Charles River Institutional Animal Care and Use Committee approved all housing and experimental procedures described (Protocol: P06282012). All procedures took place at Charles River's AAALAC accredited Kingston, NY, facility in 2012. All AAALAC accredited facilities adhere to standards set forth in the *Guide for the Care and Use of Laboratory Animals*.[26]

Animals and housing. Thirty-two Crl:CD1(ICR) (CD1) 56 to 63 d old male mice were obtained from Charles River (Kingston, NY, rooms K64 and K95. K95 has since been closed, and K64 no longer contains CD1 mice). As a direct reduction and refinement application, we chose to collect data from animals already scheduled for surgery, therefore only one sex was tested in this study. Animals were transferred to the surgical suite on-site, and singly housed 5 d prior to surgery (day -5). Each cage was comprised of a plastic bottom (29.21cm × 19.05cm × 12.70 cm; Ancare, Bellmore, NY), a wire lid, and a filtered top, as well as heat-treated hardwood bedding material (Beta Chip and aspen shavings, NEPCO, Warrensburg, NY). Animals were provided with food (Lab Diet 5L79; Purina Mills, Richmond, IN) and water without restriction. Animals were housed in a 12.5:11.5 h light-dark cycle (0430 lights on, 1700 lights off) at an average room temperature of 22 °C and humidity of 40% to 45%. All animals were monitored for a list of common mouse infectious agents that correspond roughly to the FELASA guidelines;[39] representative results may be found at https://www.criver.com/HealthData/na/H03K640M.pdf. At the time of the experiment, both rooms tested positive for *S. aureus*, but all other monitored infectious agents were negative.

Experimental Design. Power cannot be directly calculated for complex factorial designs.[11] Therefore, we followed Mead resource equation[42] to estimate the a priori sample size required to answer our experimental questions. This study used a 2 (nesting condition) × 2 (analgesia condition) × 2 (surgical condition)

factorial design with 4 replicates (*n* = 4; 32 mice total; Table 1). Mice were randomly assigned to one of the 8 experimental combinations. Mice assigned to the positive nesting treatment (+Nest) were provided with 10 grams of crinkle paper nesting material (Enviro-dri, Fibercore, Cleveland, OH) on arrival to the surgical suite. Mice assigned to the negative nesting treatment (-Nest) were not given any nesting material, however mice could make cup-shaped nests with the provided bedding material.[23] The 2 analgesia conditions included an analgesia control (-A), where mice were given a saline injection and anesthetized (43 mg/kg ketamine and 8.7 mg/kg xylazine, intraperitoneally) on the day of surgery but did not receive analgesia, and an analgesic treatment (+A), where mice received anesthesia as well as 1 subcutaneous dose of the analgesic buprenorphine on the day of surgery (0.05 mg/kg buprenorphine, Henry Schein).[12,50] The surgery conditions included anesthesia and carotid artery catheterization surgery (+Sx) and an anesthesia-only group where mice were anesthetized and placed in a surgical hood without receiving carotid artery catheterization (-Sx). Mice from each of the 8 possible experimental condition combinations were randomized into 4 observation cohorts to account for the differences in time of day or surgical order. Three animals died during surgery or recovery, 2 from the -Nest/+A/+Sx combination, and one from the +Nest/-A/+Sx combination. Only presurgical data from these animals are included in analyses.

Mice had a 72-h acclimation period before data were collected. After the first acclimation day (day -5), the mice were given 2 TINT practice sessions the following 2 mornings (days -4 and -3), since our previous work indicates that mice need exposure to the TINT for assay validity.[50] Two days before surgery (Day -2, -1: baseline; Figure 1), behavioral measurements were recorded for each animal between the hours of 0700 and 1000. First, each successive cohort was assessed for nest quality score (NQS) and grimace scores (MGS; see below for specific methods). Data from these 2 d were treated as a repeated measure for baseline. Following those observations, each cohort was tested with the TINT. After behavioral measurements were complete, body and food weights were taken.

On the day of surgery (Day 0), each cohort underwent surgical preparation and procedures concurrently. All procedures were completed between the hours of 0500 and 1200. Three hours after recovering from anesthesia, each mouse was assessed with MGS and TINT (Day 0; Post-Sx). After behavioral measurements were complete, body and food weights were taken.

One day after surgery (Day 1; Pre-Rescue), each mouse was assessed for NQS, MGS, and TINT also between 0700 and 1000. After behavioral measurements were complete, body and food weights were taken. Each animal was then given a single subcutaneous injection of buprenorphine (0.05 mg/kg, Henry Schein). One hour after injection (Day 1; Post-Rescue), all animals were reassessed using the MGS and TINT procedure and body and food weights were then collected.

Recovery measurements were taken on the following 3 days (Days 2, 3 and 4; Recovery). Each mouse was assessed for NQS, MGS, and TINT between the hours of 0700 and 1000. Data from these days was treated as a repeated measure for the recovery period. Body and food weights were then measured.

Surgical Procedure. Surgical procedures were performed by a highly trained surgeon according to Charles River's inhouse surgical protocol for carotid artery catheterization surgery. To reduce the number of animals subjected to potentially painful procedures, some of the mice tested in this experiment were to be used for other scientific purposes. As a result, we did not have full control over surgical protocols or husbandry

Table 1. This table illustrates the number of male mice in each treatment combination for this study's factorial design. Three animals were lost during surgery. Thus, the number in parentheses indicate the number of mice that data was collected from during rescue and recovery time points.

	+Sx		-Sx	
	-A	+A	-A	+A
+Nest	4 (3)	4	4	4
-Nest	4	4 (2)	4	4
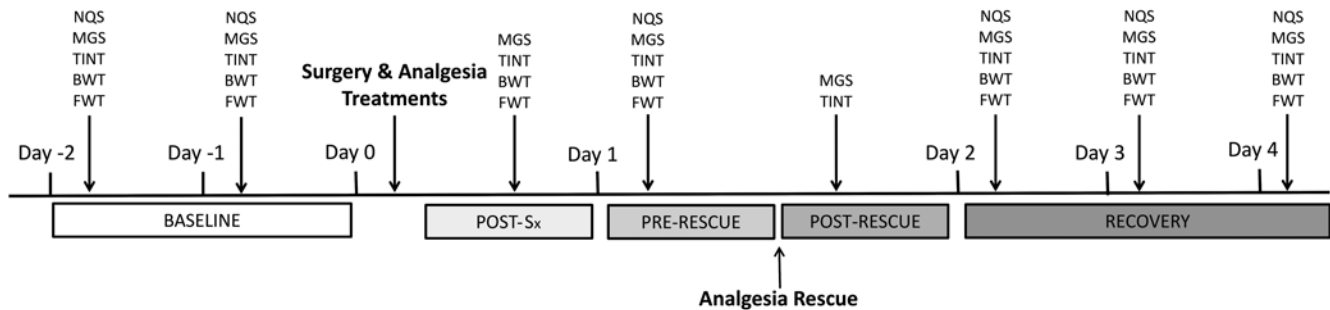
Figure 1. Timeline of experimental design. The baseline time point represents Day -2 and Day -1, or the 2 d prior to surgery or anesthesia alone. On Day 0 either surgery or anesthesia only was performed, with measurements being taken at the Post-Sx time point. On Day 1 measurements were taken before (Pre-Rescue) and after (Post-Rescue) buprenorphine administration. The recovery time point represents Days 2–4, following surgery or anesthesia only. NQS: nest scores; MGS: both live and video recorded mouse grimace scale; TINT: Time to integrate into nest test; BWT: mouse body weight; FWT: food weight.

conditions. All mice, including the anesthesia only condition, received the same presurgical preparation of the potential surgical site being shaved and disinfected. All animals were first given an intraperitoneal anesthetic injection (43 mg/kg ketamine and 8.7 mg/kg xylazine, Henry Schein). Next, they received either a subcutaneous injection of analgesic (0.05 mg/kg buprenorphine,[12,50] or vehicle injection (0.02 mL sterile saline), depending on their assigned surgical treatment. All mice had a sterile, lubricating eye ointment applied, the ventral neck and dorsal intrascapular region were shaved, and the skin was prepared for the surgery with alternating scrubs of an iodophor disinfectant and isopropyl alcohol. Stanfield heating pads (Osborne Industries, Osborne, KS), maintained at 25 to 30 °C, were used as a supplementary heating source during and after surgery for all animals. Animals assigned to the surgical cohort were transferred to the Charles River surgeon for carotid artery catheterization. A small skin incision was made on the ventral neck, exposing the left carotid artery. The catheter was inserted into the vessel, anchored with 6-0 monofilament suture, and tunneled subcutaneously, then exteriorized in the scapular region. Patency and position of the catheter was checked and then stabilized with a wound clip. The vessel incision site was closed with 5-0 monofilament suture.

Mice in the anesthesia-only condition were placed in a standard cage within a surgical hood, on a heating pad, for the same duration of time that a surgical condition mouse was operated upon. All mice recovered in a cage resting on a heating pad until they were ambulatory. Once ambulatory, the mice were returned to a clean home cage. Mice from the +N treatment also received their home nest, transferred from their previous cage. All mice had access to a heated portion of the cage overnight by placing a portion of the cage on a heating pad.

Measures. Attempts to blind researchers or husbandry staff to nest and surgical treatments could not be achieved during live assessments because of the obvious visual presence of these treatments.

TINT procedure. The TINT procedure followed the protocol published by Rock and colleagues[50] All TINT observations were made between the hours of 0700 and 1000, unless otherwise indicated in the study design. The cage lid was lifted and a 1"x 1" compressed cotton square (Cotton square, Ancare, Bellmore, NY) was dropped into the side of the cage opposite the nest. Latencies to interact with the TINT material into the existing nest were recorded. If the nesting material was incorporated into the existing nest or sleeping site within 10 min of its introduction, the mouse was assigned a positive TINT. If the TINT material was not incorporated or incorporated after 10 min, the mouse was assigned a negative TINT. TINT was assessed by a female observer (MSG).

Nest Quality Scoring (NQS). The nest from each cage was scored using a naturalistic 0 to 5-point nesting scale.[16,24] A score of 0 to 5 was given to each of the 4 sides of the nest and were averaged to arrive on the cumulative nest quality score. Briefly, a score of 0: indicated that the material had not been manipulated; 1: indicated it had been manipulated but no clear nest site was present; 2: a nest site was present but the nest was flat; 3: a shallow cup shape was present; 4: the wall of the nest was less than half the height of a dome; and 5: the wall was greater than half the height of the dome which may or may not fully enclose the nest. The live female scorer (MSG) was trained by BNG, the developer of the scale, before collecting data.

Mouse Grimace Scale (MGS) Scoring. Mice were assessed using the scale developed by Langford and colleagues[35] Each cage was removed from the rack and placed on a table for cage-side observation by a single female observer (MSG). The observer was trained with still images to identify MGS prior to data collection with the methods published by Langford and colleagues.[35] In this experiment, we assessed the MGS live as well as from images produced from a 30 s long videos recorded at the same time as the live, cage side, assessment. The mouse was assessed only when outside of the nest and when determined to be fully alert. The scorer observed each mouse for at least 30 s and assessed it for orbital tightening, ear position, whisker change, cheek bulge, and nose bulge on a 3-point scale (0= None, 1= Moderate, 2= Severe). The videos were later processed through iMovie 11 for Mac (version 9.0.4, Apple), where individual frames were exported from the video for still images. Next the exported images were evaluated for

quality (photos were omitted if blurry or if the mouse was not facing the camera). Of the remaining images, 2 per mouse per time point were randomly chosen (www.random.org; random number generator) for scoring. The images were pasted into a PowerPoint file and were cropped so that only the mouse's face was visible. Images were distributed between 2 observers blinded to the assigned treatments for evaluation (BNG and SYD). Both scorers were trained over several weeks prior to scoring MGS images. If a scorer did not feel the image quality was good enough to confidently evaluate a specific action unit (such as whisker position) that score was left blank. Action units were averaged over the 2 separate images per mouse, per time point. Due to lost data in the image MGS assessment, only data from day 2 was used as the recovery time point comparisons between the MGS live score, MGS image score, and TINT.

Food Consumption and Body Weight Monitoring. All cages were monitored daily for the animal body and food weights after experimental manipulations. The weight of the food was documented at the start of the study and weighed daily. If food was added to the hopper, it was weighed. Animal weight change was calculated as the difference in body weight from the mouse's weight the previous day, following how animals gained or lost weight over the experiment.

Statistical Procedures. All statistical analyses were performed using JMP v.10 statistical (SAS Institute, Cary, NC). Repeated TINT outcomes were analyzed as a binary logistic regression with the generalized linear model (GLIM) using the Firth adjusted maximum likelihood estimation method. To avoid pseudoreplication and accommodate repeated measures, analyses were blocked by cage, nested within surgical treatment, analgesic treatment, and nesting treatment. All second order interactions between surgical treatment, analgesic treatment, nesting treatment, and time point as well as the third order interaction between surgical treatment, analgesic treatment, and nesting treatment were included in the model. Post-hoc planned contrasts were used to compare significant effects and Bonferroni correction was applied for multiple comparisons. Analyses for all nonbinary dependent variables (weight change, time to interact, NQS, MGS, and food consumption) were performed using a general linear model (GLM). The assumptions of GLM (normality of error, homogeneity of variance, and linearity) were confirmed graphically post hoc.[18] The analyses were blocked by the individual, nested within surgical condition, analgesic condition, and nesting condition. Since animals were singly housed, cage was treated as a random effect. Cohort, rack side, and rack shelf blocking factors were tested to determine if they significantly altered results but were not significant and were therefore not included in any analyses. Explanatory variables found to be significant were analyzed using post hoc Tukey pairwise comparisons or with Bonferroni corrected test slices. Only the time to interact data required logarithmic transformation to achieve normality. All data are presented as least squares means ± SEs (LSM ± SE).

To determine the convergent validity between TINT, live MGS, and image MGS a principal component analysis (PCA) with varimax rotation was performed in JMP v12. Since the measures were expected to change even within one mouse over the course of the study, data from only one time point (Day 1; Pre-Rescue) was analyzed to maximize variability in the data. Only principal components with an eigenvalue greater than 1 were evaluated.

Results

Analysis details are summarized in Table 2 for the results reported below. Post hoc test slices or planned contrasts are reported in full detail unless a Tukey test was run. Significant Tukey tests are indicated by $P < 0.05$.

TINT. Mice in the -Nest condition were more likely to fail the TINT throughout the experiment than those mice who were provided with nesting material ($P = 0.049$; Figure 2). The main effect of analgesia treatment (treatment assigned to each mouse, regardless of time point) did not affect the TINT ($P > 0.99$). However, the probability of failing the TINT was also affected by an interaction of time point within the study and surgical treatment ($P = 0.007$; Figure 3). Three hours after recovery from anesthesia (Day 0), mice were much more likely to fail the TINT if they had undergone surgery rather than anesthesia alone ($F_{1,194} = 7.44$; $P = 0.006$). Custom contrasts at the baseline and Pre-Rescue time points could not be calculated due to the lack of variability in the data, because all the mice had similar TINT outcomes at these time points, regardless of surgical condition (see Table 3). However, based on a post hoc χ^2 test, +Sx animals are more likely to fail at the Pre-Rescue time point than –Sx mice ($\chi2 = 16.27$; $P < 0.001$).

Time to Interact with TINT material. There was a significant difference in the time it took mice to interact with the TINT material due to the experimental time point ($P = 0.008$). At the recovery time point, mice were slower to interact with the TINT material across all treatments than they had been at baseline ($P < 0.05$). Time to interact with TINT material was also affected by a significant interaction between nesting condition, analgesic condition, and surgical condition ($P = 0.015$). Mice that did not receive buprenorphine or undergo surgery were faster to interact with the TINT material if they were assigned to the +Nest condition ($P < 0.05$).

Nest Quality Score. Mice in the +Nest condition earned a significantly higher NQS than mice not given nesting material ($P < 0.001$), despite bedding allowing a nest to be built where typically, other bedding does not. Mice in the +Sx condition had a significantly lower NQS than mice that did not undergo surgery ($P = 0.03$; Figure 4). NQS was also altered due to the experimental time point ($P < 0.001$). Mice scored the highest NQS the morning of Day 0, prior to undergoing surgery (4.30 ± 0.12) and the lowest NQS the morning of Day 1; 15 to 22 h after surgery (3.07 ± 0.12; $P < 0.05$). Baseline NQS values (3.91 ± 0.09) were different than those found on Day 0 and Day 1 but were statistically the same as recovery scores (4.07 ± 0.09). NQS values at recovery (4.06 ± 0.08) were also statistically different from Day 1 scores ($P < 0.05$). Analgesia did not significantly alter NQS ($P = 0.95$).

Mouse Grimace Score. *Live scoring.* Like the TINT results, only an interaction of surgery treatment by time point was significant ($P < 0.001$). Surgical control mice (-Sx) did not differ in their MGS scores at any time point ($P > 0.05$). Mice that received the +Sx treatment had significantly higher MGS scores directly after surgery (0.37 ± 0.037), the day after surgery at the Pre-Rescue (0.29 ± 0.038) and Post-Rescue (0.18 ± 0.038) time points, compared with scores from Baseline (0.013 ± 0.035) and Recovery (0.05 ± 0.040; all Ps < 0.05). However, Post-Rescue scores were not significantly lower than Pre-Rescue scores ($P > 0.05$) but were lower than Post-Sx scores the day before ($P < 0.05$).

Image scoring. The interaction of nesting treatment by analgesia was significant ($P = 0.042$). However, post hoc tests did not reveal any significant differences between treatments ($P > 0.05$). An interaction between surgery, analgesia, and time point was also found ($P = 0.027$). Only one post hoc comparison was found between –Sx (1.2 ± 0.13) and +Sx (0.45 ± 0.13) mice who all received analgesia, at the postsurgical time point ($P < 0.05$).

Table 2. Statistical test details from experimental treatment variables (surgical or analgesia treatments) and other significant terms in the model performed on measures investigated in this study.

Measure	Analysis	Reported Effect	DF	Denominator DF	Test statistic	P value
TINT	BLR	Nest	1		3.89	0.048
		Time Point×Surgery	4		13.98	0.0074
		Analgesia	1		<0.01	>0.99
Time to Interact	GLM	Time Point	4	164.5	3.53	0.0085
		Nest×Surgery× Analgesia	1	22.61	6.91	0.015
NQS	GLM	Nest	1	27.97	37.81	<0.001
		Time Point	3	169.9	25.06	<0.001
		Surgery	1	27.97	5.03	0.033
		Analgesia	1	27.97	0.0034	0.95
MGS Live	GLM	Surgery×Time Point	4	92.24	7.54	<0.001
MGS Image	GLM	Analgesia×Nest	1	24.01	4.63	0.042
		Analgesia× Surgery×Time Point	4	97.81	2.87	0.027
Body Weight Change	GLM	Time Point×Surgery	3	169.6	8.29	<0.001
		Analgesia	1	87.21	0.77	0.38
Food Consumption	GLM	Time Point×Surgery	3	158.6	3.49	0.017
		Analgesia	1	19	3.15	0.092

Analysis refers to the statistical test run on that measure; BLR = Binary Logistic Regression; GLM = General linear model; DF = numerator degrees of freedom; Denominator DF = denominator degrees of freedom; Test statistic = likelihood ratio χ^2 for BLR analysis and F value for GLM.

Convergent Validity of Pain Measures. The PCA analysis on the correlations of live MGS, image MGS, and TINT resulted in one component with an eigenvalue of 2.07 and 2 others with
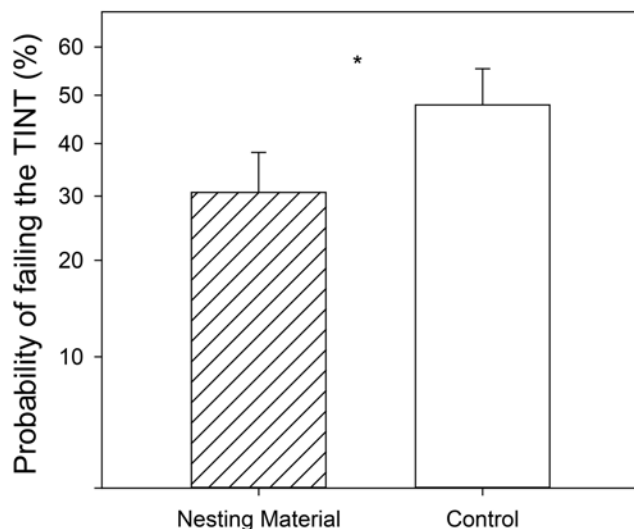

Figure 2. Effect of the main effect of nesting treatment on TINT performance. The probability of failing the TINT is represented on the y-axis on a back transformed $\log_{10}$ scale. Essentially the data graphed are the $\log_{10}$ LSM ± SE values from the analysis but only the tick marks on the graph have been back transformed to depict values in the real world. Asterisks indicate the significant effect ($P < 0.05$).

values less than 1 that did not receive further evaluation. All 3 pain measures loaded onto the first component with loadings of live MGS 0.820; image MGS 0.782; and TINT -0.890.

Mouse Weight Change. Mouse weight change was affected by an interaction between surgical condition and time point ($P < 0.001$; Figure 5). On the day after surgery (Day 1), mice in the +Sx condition lost significantly more weight (-2.32g ± 0.18) than those that were only anesthetized (-0.99g ± 0.17), when compared with their previous day's weights ($F_{1, 177.1} = 30.58$; $P < 0.001$). Much like the MGS image scores, analgesia did not alter results ($P = 0.38$).

Food Consumption. A significant interaction between time point and surgical condition affected food consumption ($P = 0.017$). Mice who received the catheterization surgery ate less food the day after surgery (4.6g ± 0.51), than –Sx mice (6.8g ± 0.48; Day 1; $F_{1,150} = 9.33$, $P = 0.0027$). Analgesic treatments however did not affect food consumption ($P = 0.092$).

Discussion

The goal of this study was to validate whether the TINT could correctly identify mice that were in pain. The TINT showed discriminant validity by differentiating between mice that had undergone surgery compared with those that were simply anesthetized. Animals receiving surgery were significantly more likely to fail the TINT than sham-operated controls. However, the drugs given to animals on the established surgical protocol inhibited our ability to interpret the analgesia condition, preventing us from identifying protection from or reversal of postoperative pain. Since conducting this project in 2012, many publications have shown that the recommended doses of buprenorphine[12] are either partially or not completely effective.[28,32,47] Therefore it may not be surprising that we did not find significant effects of analgesic treatment. However, almost all mice, regardless of surgical or analgesia treatments, failed the TINT after receiving the rescue analgesia dose the day after surgery. This finding provides further data indicating that the TINT is altered under many experimental conditions and as such, we were not able to demonstrate discriminant validity by reversing the TINT with the analgesia used in this experiment.
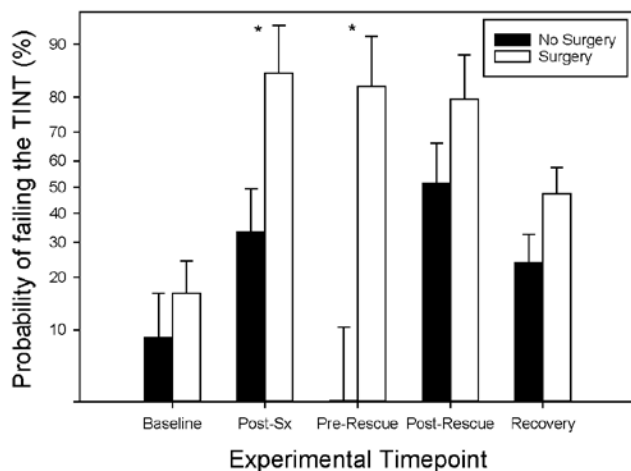
We had expected the most painful experimental time points to be at Post-surgery and Pre-Rescue, resulting in an expected increase in the likelihood of TINT failure through disrupted nesting behavior. The rescue element of the experiment was designed

Figure 3. Effect of experimental time point and surgical condition on performance on the TINT. The probability of failing the TINT is represented on the y-axis on a back transformed $\log_{10}$ scale. Asterisks indicate the significant difference between –Sx and +Sx groups at the Post-Sx and Pre-Rescue time points ($P < 0.05$). Baseline data is from days -1 and -2 while data from the recovery time point includes days 2, 3, and 4. Data is presented as LSM ± SE.

to evaluate if mice that fail the TINT initially could reverse their result to a positive TINT after analgesic administration, indicating that the TINT failure was due to the influence of pain on behavior. Buprenorphine is known to have several behavioral side effects that are inextricable from its analgesic properties.[5,13,22] However, due to its effectiveness, buprenorphine is used as the primary postoperative analgesic for Charles River's surgical procedures. We attempted to embrace the 3Rs by using animals already scheduled for routine surgery, therefore the use of buprenorphine as the chosen analgesic was predetermined, despite its limitations in a behavioral study. During the Pre-Rescue time point, pain did indeed cause mice to fail the TINT, since most all –Sx mice passed and +Sx mice failed when no analgesics or anesthetic drugs were present. However, after the rescue buprenorphine was administered (Post-Rescue), when +Sx mice should have been experiencing pain relief, both +Sx and –Sx mice failed the TINT. Failure of both surgical and control groups implicates buprenorphine in the disruption of the mice's ability to perform the TINT at this time point. Whether pain was relieved in the +Sx mice is unknown due to the drug's known effects on behavior.

Like other opioid analgesics, buprenorphine is known to cause behavioral and physiologic side effects in many species even at clinically relevant analgesic doses.[19] These nonspecific drug effects could potentially alter the drug's analgesic efficacy. Our findings indicate that nest building behavior as measured by the TINT is susceptible to alteration by the non-analgesic effects of buprenorphine. This variable potential for behavioral changes, not related to pain, to occur with commonly used analgesics underscores the frustrating challenge confronting researchers attempting to develop pain scales without interference from nonspecific drug effects.

Another potential confounding influence on TINT outcomes could be the anesthetic protocol. The ketamine and xylazine mixture used to anesthetize mice for surgery/sham treatment may have provided a moderate analgesic benefit to every mouse in the study.[44,48] We had hoped providing a level of analgesia in the anesthetic protocol would mitigate some of the pain experienced by our +Sx/−A mice. However, this protocol may have confounded our analgesia condition. Future studies evaluating the possibility of TINT to identify postoperative pain in mice should use inhalant anesthetics such as sevoflurane or isoflurane

which have no residual analgesic effect.[7] By eliminating the confounding pharmacologic effects of ketamine and xylazine, the protective effects of a chosen analgesic could be evaluated in isolation.

When developing new assessment protocols, the technique must be shown to measure a meaningful or valid condition. To demonstrate convergent construct validity, we included MGS, body weight, and food consumption in this study. Although these measures are commonly used to assess pain, body weight and food consumption are general measures of welfare, but are also affected by analgesic treatment.[41] The MGS has succeeded in assessing postoperative analgesic treatments irrespective of the drug class being evaluated, identifying pain, and evaluating the efficacy of commonly used analgesics.[35,36,41] The MGS is clearly the most specific and valid measure of pain in our study, making it an important standard to assess the TINT against.

The PCA resulted in one component in which the image MGS, live MGS, and TINT all had similar strong loading values, indicating that all 3 measures are highly intercorrelated. The negative correlation of TINT compared with either MGS scoring methods (live or via photos) corresponds to animals in more pain having an increased MGS score and decreased TINT values. However, one study recently reported significant differences between live and image MGS scores.[43] Although all 3 of our measures showed convergent construct validity, we recognize the potential inherent bias of live assessment and the difficulty in observing subtle changes in quickly moving mice. The TINT was originally developed as a tool for quick cage-side assessment of pain; however, if live, unblinded, scoring is biased, we assume it might produce more false positives (mice identified as painful, even if they're not). As long as a surgery isn't part of an experimental treatment, this is likely to be better for animal welfare since providing analgesia to a mouse that doesn't need it is preferred to denying it to a mouse that does.

In this study, the MGS image assessment was challenging. Our images were extracted from video frames and were not of sufficient quality. This caused us to omit whisker position and nose bulge quite frequently. Image extraction to evaluate MGS is only recommended if a higher speed, professional-grade camera is used.

Each experimental manipulation in our study: nesting material, surgery, and analgesia affected the metabolic needs of the animal. Postsurgical catabolism is a well-recognized phenomenon in all vertebrates; the stress response affects metabolism and depresses feed intake.[10,40] Buprenorphine is also known to decrease food intake and growth rates[5,6,25,27] and similarly, nesting material provides significant thermal benefit resulting in improved calorie conversion and thus, decreased food intake.[15] While food consumption and body weight were not themselves correlated with TINT performance, they were both negatively affected 24 h after surgery. This highlights the time scales of different general pain measures. The capacity for real time evaluation of distress in mice is a strong potential benefit of the TINT not shared by the other physiologic measurements often used to monitor mouse wellbeing.

This study was designed around the assumption that nesting behavior can be used as an indicator of wellbeing in mice. Indeed, impaired nest quality has been associated with rodent postoperative pain.[3,31] As anticipated, mice in the +Sx condition had poorer nesting performance than their –Sx counterparts. Furthermore, all mice made less complex nests on Day 1, presumably due to a disruption of normal behavioral time budgets by the anesthesia administered to all mice. This validates our supposition that changes in nesting behavior reflects

Table 3. TINT Performance: the frequency of TINT successes and failures across surgical and analgesia treatments. Baseline data contains 2 d of TINT (Day -2 and -1). Recovery data contains tests from days 2, 3 and 4.

	-Sx				+Sx			
	-A		+A		-A		+A	
Time point	- TINT	+ TINT	- TINT	+ TINT	- TINT	+ TINT	- TINT	+ TINT
Baseline	1	15	3	13	1	15	4	12
Post-Sx	1	7	4	4	7	0	6	2
Pre-Rescue	0	8	1	7	7	0	4	3
Post-Rescue	3	5	4	4	5	2	6	1
Recovery	5	19	7	17	9	12	7	11
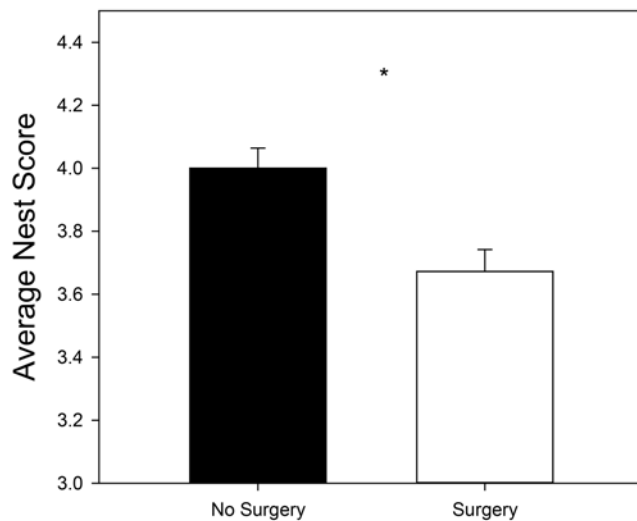
Figure 4. Effect of surgical condition on nest quality score. Mice who experienced anesthesia alone (-Sx) had significantly higher NQSs than those mice who received surgery (+Sx). Asterisk indicates the significant effect ($P < 0.05$). Data is presented as LSM ± SE.
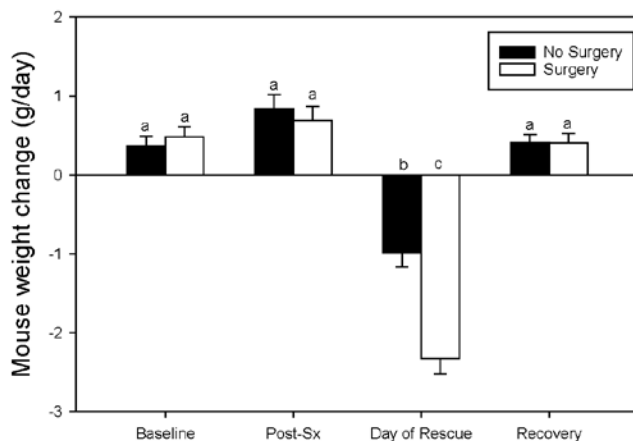
Figure 5. Effect of time point and surgical condition on change in mouse body weight from the previous day. Different letters over the bars indicate significant differences between pairwise Tukey tests ($P < 0.05$), thus if 2 letters are the same there is no statistical difference between weight change LSM values. Baseline data is from days -1 and -2 while recovery data includes days 2, 3, and 4. Data is presented as LSM ± SE.

an alteration in normal mouse behavior as a whole, whether due to pain or other influences. Nest scoring still requires time for the animals to engage in a particular sequence of behaviors before differences can be detected. These drawbacks illustrate the utility of the TINT in time needed for assessment (that is 10 min compared with 24 h) and the ability to get an instantaneous snapshot of the mouse's welfare.

Due to its novelty, some investigators have used the TINT to evaluate pain or discomfort in mice without proper validation or scientific evaluation.[21,23,34,45,47,51,52] While used in several experimental studies as a measure of pain or general discomfort, all but 2 diverged from the originally published protocol.[21,23] While it may be tempting to fit a new protocol to individual needs, this has the possibility of negating earlier validation work and could possibly give a slightly altered outcome that affects reproducibility. The TINT was initially developed to address an unmet need for a quick, objective, and reliable cage-side measure of pain in laboratory mice. Our results suggest that the TINT can identify mice that have undergone surgery and other sources of distress in mice, but that disruption of the TINT is not necessarily specific to pain resulting from a surgery. A generalized welfare indicator has far more utility than does one specific to a certain state. Possible uses can now extend to any research scenario that asks the question "Is a mouse's normal behavior disrupted?" This could include identifying changes in behavioral phenotypes of new transgenic models, nonspecific drug effects of novel therapeutic compounds, or humane endpoints in disease models. The TINT requires further validation using different analgesics, but preliminary data is encouraging that the TINT will be a valuable tool in the real-time assessment of mouse welfare.

References

1. **Abelson KSP, Jacobsen KR, Sundbom R, Kalliokoski O, Hau J.** 2012. Voluntary ingestion of nut paste for administration of buprenorphine in rats and mice. Lab Anim **46:**349–351. https://doi.org/10.1258/la.2012.012028.
2. **Anil SS, Anil L, Deen J.** 2002. Challenges of pain assessment in domestic animals. J Am Vet Med Assoc **220:**313–319. https://doi.org/10.2460/javma.2002.220.313.
3. **Arras M, Rettich A, Cinelli P, Kasermann HP, Burki K.** 2007. Assessment of post-laparotomy pain in laboratory mice by telemetric recording of heart rate and heart rate variability. BMC Vet Res **3:**1–10. https://doi.org/10.1186/1746-6148-3-16.
4. **Aubert A, Goodall G, Dantzer R, Gheusi G.** 1997. Differential effects of lipopolysaccharide on pup retrieving and nest building in lactating mice. Brain Behav Immun **11:**107–118. https://doi.org/10.1006/brbi.1997.0485.
5. **Bourque SL, Adams MA, Nakatsu K, Winterborn A.** 2010. Comparison of buprenorphine and meloxicam for postsurgical analgesia in rats: effects on body weight, locomotor activity, and hemodynamic parameters. J Assoc Lab Anim Sci **49:**617–622.

6. **Brennan MP, Sinusas AJ, Horvath TL, Collins JG, Harding MJ.** 2009. Correlation between body weight changes and postoperative pain in rats treated with meloxicam or buprenorphine. Lab Anim (NY) **38:**87–93. https://doi.org/10.1038/laban0309-87.

7. **Brunson DB.** 2008. Pharmacology of inhalation anesthetics. p 83–85. In: Fish R, Danneman P, Brown M, Karas AZ, editors. Anesthesia and analgesia in laboratory animals. Burlington (MA): Academic Press. https://doi.org/10.1016/B978-012373898-1.50007-3

8. **Carstens E, Moberg GP.** 2000. Recognizing pain and distress in laboratory animals. ILAR J **41:**62–71. https://doi.org/10.1093/ilar.41.2.62.

9. **Colpaert FC, Tarayre JP, Alliaga M, Bruins Slot LA, Attal N, Koek W.** 2001. Opiate self-administration as a measure of chronic nociceptive pain in arthritic rats. Pain **91:**33–45. https://doi.org/10.1016/S0304-3959(00)00413-9.

10. **Cowan A, Doxey J, Harry E.** 1977. The animal pharmacology of buprenorphine, an oripavine analgesic agent. Br J Pharmacol **60:**547–554. https://doi.org/10.1111/j.1476-5381.1977.tb07533.x.

11. **Festing MF, Overend P, Gaines Das R, Cortina-Borja M, Berdoy M.** 2002. The design of animal experiments: reducing the use of animals in research through better experimental design. United Kingdom. The Royal Society of Medicine Press Limited.

12. **Flecknell P.** 2009. Laboratory animal anaesthesia. London (United Kingdom): Academic Press.

13. **Gades NM, Wixson SK, Danneman PJ, Tolley EA.** 2001. Inefficacy of buprrenorphine in rats. Contemp Top Lab Anim Sci **40:**7.

14. **Gaskill BN, Gordon CJ, Pajor EA, Lucas JR, Davis JK, Garner JP.** 2012. Heat or insulation: Behavioral titration of mouse preference for warmth or access to a nest. PLoS One **7:**1–11. https://doi.org/10.1371/journal.pone.0032799.

15. **Gaskill BN, Gordon CJ, Pajor EA, Lucas JR, Davis JK, Garner JP.** 2013. Impact of nesting material on mouse body temperature and physiology. Physiol Behav **110–111:**87–95. https://doi.org/10.1016/j.physbeh.2012.12.018.

16. **Gaskill BN, Pritchett-Corning KR.** 2016. Nest building as an indicator of illness in laboratory mice. Appl Anim Behav Sci **180:**140–146. https://doi.org/10.1016/j.applanim.2016.04.008.

17. **Gaskill BN, Pritchett-Corning KR, Gordon CJ, Pajor EA, Lucas JR, Davis JK, Garner JP.** 2013. Energy reallocation to breeding performance through improved nest building in laboratory mice. PLoS One **8:**1–9. https://doi.org/10.1371/journal.pone.0074153.

18. **Grafen A, Hails R.** 2002. Modern statistics for the life sciences. Oxford; New York, Oxford University Press.

19. **Grimm KA, Lamont LA, Tranquilli WJ, Greene SA, Robertson SA, editors.** 2015. Veterinary anesthesia and analgesia. Ames (IA): Wiley–Blackwell.

20. **Gross AN-M, Engel AKJ, Würbel H.** 2011. Simply a nest? Effects of different enrichments on stereotypic and anxiety-related behaviour in mice. Appl Anim Behav Sci **134:**239–245. https://doi.org/10.1016/j.applanim.2011.06.020.

21. **Häger C, Keubler LM, Biernot S, Dietrich J, Buchheister S, Buettner M, Bleich A.** 2015. Time to integrate to nest test evaluation in a mouse DSS-colitis model. PLoS One **10:**1–12. https://doi.org/10.1371/journal.pone.0143824.

22. **Hayes LD.** 2000. To nest communally or not to nest communally: a review of rodent communal nesting and nursing. Anim Behav **59:**677–688. https://doi.org/10.1006/anbe.1999.1390.

23. **Herndon NL, Bandyopadhyay S, Hod EA, Prestia KA.** 2016. Sustained-release buprenorphine improves postsurgical clinical condition but does not alter survival or cytokine levels in a murine model of polymicrobial sepsis. Comp Med **66:**455–462.

24. **Hess SE, Rohr S, Dufour BD, Gaskill BN, Pajor EA, Garner JP.** 2008. Home improvement: C57BL/6J mice given more naturalistic nesting materials make better nests. J Am Assoc Lab Anim Sci **47:**25–31.

25. **Ilbäck NG, Siller M, Stålhandske T.** 2008. Effects of buprenorphine on body temperature, locomotor activity and cardiovascular function when assessed by telemetric monitoring in rats. Lab Anim **42:**149–160. https://doi.org/10.1258/la.2007.06002e.

26. **Institute for Laboratory Animal Research.** 2011. Guide for the care and use of laboratory animals. Washington (DC): National Academies Press.

27. **Jablonski P, Howden BO, Baxter K.** 2001. Influence of buprenorphine analgesia on postoperative recovery in 2 strains of rats. Lab Anim **35:**213–222. https://doi.org/10.1258/0023677011911651.

28. **Jirkof P.** 2014. Burrowing and nest building behavior as indicators of wellbeing in mice. J Neurosci Methods **234:**139–146. https://doi.org/10.1016/j.jneumeth.2014.02.001.

29. **Jirkof P.** 2017. Side effects of pain and analgesia in animal experimentation. Lab Anim (NY) **46:**123–128. https://doi.org/10.1038/laban.1216.

30. **Jirkof P, Cesarovic N, Rettich A, Nicholls F, Seifert B, Arras M.** 2010. Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci **4:**1–9. https://doi.org/10.3389/fnbeh.2010.00165.

31. **Jirkof P, Fleischmann T, Cesarovic N, Rettich A, Vogel J, Arras M.** 2013. Assessment of postsurgical distress and pain in laboratory mice by nest complexity scoring. Lab Anim **47:**153–161. https://doi.org/10.1177/0023677213475603.

32. **Jirkof P, Tourvieille A, Cinelli P, Arras M.** 2014. Buprenorphine for pain relief in mice: repeated injections vs sustained-release depot formulation. Lab Anim **49:**177–187. https://doi.org/10.1177/0023677214562849.

33. **Karas AZ.** 2006. Barriers to assessment and treatment of pain in laboratory animals. Lab Anim (NY) **35:**38–45. https://doi.org/10.1038/laban0706-38.

34. **Kendall LV, Wegenast DJ, Smith BJ, Dorsey KM, Kang S, Lee NY, Hess AM.** 2016. Efficacy of sustained-release buprenorphine in an experimental laparotomy model in female mice. J Am Assoc Lab Anim Sci **55:**66–73.

35. **Langford DJ, Bailey AL, Chanda ML, Clarke SE, Drummond TE, Echols S, Glick S, Ingrao J, Klassen-Ross T, LaCroix-Fralish ML, Matsumiya L, Sorge RE, Sotocinal SG, Tabaka JM, Wong D, van den Maagdenberg A, Ferrari MD, Craig KD, Mogil JS.** 2010. Coding of facial expressions of pain in the laboratory mouse. Nat Methods **7:**447–449. https://doi.org/10.1038/nmeth.1455.

36. **Leach MC, Klaus K, Miller AL, di Perrotolo MS, Sotocinal SG, Flecknell PA.** 2012. The assessment of post-vasectomy pain in mice using behaviour and the Mouse Grimace Scale. PLoS One **7:**1–9.

37. **Lee CT.** 1972. Development of nest-building behavior in inbred mice. J Gen Psychol **87:**13.

38. **Littin K, Acevedo A, Browne W, Edgar J, Mendl M, Owen D, Sherwin C, Wurbel H, Nicol C.** 2008. Towards humane end points: behavioural changes precede clinical signs of disease in a Huntington's disease model. Proc R Soc Biol Sci Ser B **275:**1865–1874. https://doi.org/10.1098/rspb.2008.0388

39. **Mähler Convenor_M, Berard M, Feinstein R, Gallagher A, Illgen-Wilcke B, Pritchett-Corning K, Raspa M.** 2014. FELASA recommendations for the health monitoring of mouse, rat, hamster, guinea pig and rabbit colonies in breeding and experimental units. Lab Anim **48:**178–192. https://doi.org/10.1177/0023677213516312. Erratum. [Lab Anim 2015].

40. **Marquez P, Baliram R, Kieffer BL, Lutfy K.** 2007. The mu opioid receptor is involved in buprenorphine-induced locomotor stimulation and conditioned place preference. Neuropharmacology **52:**1336–1341. https://doi.org/10.1016/j.neuropharm.2007.01.005.

41. **Matsumiya LC, Sorge RE, Sotocinal SG, Tabaka JM, Wieskopf JS, Zaloum A, King OD, Mogil JS.** 2012. Using the mouse grimace scale to reevaluate the efficacy of postoperative analgesics in laboratory mice. J Am Assoc Lab Anim Sci **51:**42–49.

42. **Mead R.** 1988. The design of experiments: statistical principles for practical applications. New York (NY): Cambridge University Press.

43. **Miller AL, Leach MC.** 2015. The Mouse Grimace Scale: A clinically useful tool? PLoS One **10:**1–10. https://doi.org/10.1371/journal.pone.0136000.

44. **Muir WW.** 1985. Cyclohexanone drug mixtures: The pharmacology of ketamine and ketamine drug combinations, Proc 2nd Intl Cong Vet Anes. 5–14.

45. **Negus SS, Neddenriep B, Altarifi AA, Carroll FI, Leitl MD, Miller LL.** 2015. Effects of ketoprofen, morphine, and kappa opioids on pain-related depression of nesting in mice. Pain **156:**1153–1160.
46. **Nijsen MJ, Ongenae NG, Coulie B, Meulemans AL.** 2003. Telemetric animal model to evaluate visceral pain in the freely moving rat. Pain **105:**115–123. https://doi.org/10.1016/S0304-3959(03)00170-2.
47. **Oliver VL, Thurston SE, Lofgren JL.** 2018. Using cageside measures to evaluate analgesic efficacy in mice (*Mus musculus*) after surgery. J Am Assoc Lab Anim Sci **57:**186–201.
48. **Persson J.** 2013. Ketamine in pain management. CNS Neurosci Ther **19:**396–402. https://doi.org/10.1111/cns.12111.
49. **Rock ML, Karas AZ, Gallo MS, Pritchett-Corning KR, Gaskill BN.** 2014. Housing condition and nesting experience do not affect the Time to Integrate to Nest Test (TINT). Anim Welf **23:**381–385. https://doi.org/10.7120/09627286.23.4.381.
50. **Rock ML, Karas AZ, Rodriguez KBG, Gallo M, Pritchett-Corning KR, Gaskill BN.** 2014. The time-to-integrate-to-nest test as an indicator of wellbeing in laboratory mice. J Am Assoc Lab Anim Sci **53:**24–28.
51. **Toma W, Kyte SL, Bagdas D, Alkhlaif Y, Alsharari SD, Lichtman AH, Chen Z-J, Del Fabbro E, Bigbee JW, Gewirtz DA.** 2017. Effects of paclitaxel on the development of neuropathy and affective behaviors in the mouse. Neuropharmacology **117:**305–315. https://doi.org/10.1016/j.neuropharm.2017.02.020.
52. **Yuan D, Liu C, Wu J, Hu B.** 2018. Nest-building activity as a reproducible and long-term stroke deficit test in a mouse model of stroke. Brain Behav **8:**1–9. https://doi.org/10.1002/brb3.993.