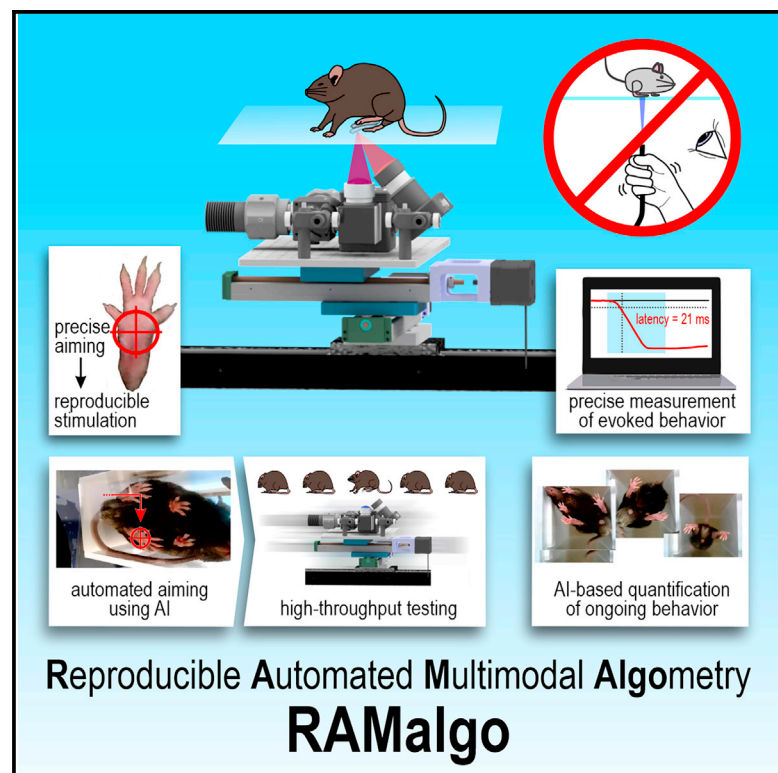


Reproducible and fully automated testing of nocifensive behavior in mice

Graphical abstract



Authors

Christopher Dedek, Mehdi A. Azadgoleh, Steven A. Prescott

Correspondence

steve.prescott@sickkids.ca

In brief

Preclinical pain testing traditionally utilizes techniques—such as stimulating by hand and measuring responses by eye—that limit reproducibility. Dedek et al. describe a device for consistent stimulation and withdrawal measurement with millisecond precision. Testing can be fully automated, increasing standardization and throughput while objectively quantifying non-evoked behaviors.

Highlights

- Poor reproducibility has plagued preclinical pain research and demands improvement
- We present a technology to standardize stimulation and precisely measure withdrawal
- Machine learning automates aiming and quantification of evoked and non-evoked behaviors
- Automation improves reproducibility, throughput, and comprehensiveness of pain testing



Article

Reproducible and fully automated testing of nocifensive behavior in mice

Christopher Dedek,^{1,2} Mehdi A. Azadgoleh,¹ and Steven A. Prescott^{1,2,3,4,*}¹Neurosciences and Mental Health, The Hospital for Sick Children, Toronto, ON M5G 0A4, Canada²Institute of Biomedical Engineering, University of Toronto, Toronto, ON M5S 3G9, Canada³Department of Physiology, University of Toronto, Toronto, ON M5S 1A8, Canada⁴Lead contact*Correspondence: steve.prescott@sickkids.ca<https://doi.org/10.1016/j.crmeth.2023.100650>

MOTIVATION Preclinical pain testing suffers from poor reproducibility, low throughput, and a restricted focus on reflex responses. To address the first two problems, we sought to replace manual stimulation and response measurement with a pain-testing robot. To address the third problem, the robot collects video amenable to the latest machine-learning-based analysis tools to quantify non-reflexive behaviors. Last, to facilitate creation of large datasets, we developed software to store and organize collected data.

SUMMARY

Pain in rodents is often inferred from their withdrawal from noxious stimulation. Threshold stimulus intensity or response latency is used to quantify pain sensitivity. This usually involves applying stimuli by hand and measuring responses by eye, which limits reproducibility and throughput. We describe a device that standardizes and automates pain testing by providing computer-controlled aiming, stimulation, and response measurement. Optogenetic and thermal stimuli are applied using blue and infrared light, respectively. Precise mechanical stimulation is also demonstrated. Reflectance of red light is used to measure paw withdrawal with millisecond precision. We show that consistent stimulus delivery is crucial for resolving stimulus-dependent variations in withdrawal and for testing with sustained stimuli. Moreover, substage video reveals “spontaneous” behaviors for consideration alongside withdrawal metrics to better assess the pain experience. The entire process was automated using machine learning. RAMalgo (reproducible automated multimodal algorithm) improves the standardization, comprehensiveness, and throughput of preclinical pain testing.

INTRODUCTION

Measuring withdrawal from noxious stimuli in laboratory rodents is a mainstay of preclinical pain research.^{1–4} Testing is often conducted on the hind paw, in part because many chronic pain models are designed to increase paw sensitivity through manipulations of the paw or the nerves innervating it.^{4–7} Measuring evoked pain with withdrawal reflexes has been criticized⁸ because ongoing (non-evoked) pain is a bigger clinical problem,⁹ but tactile and thermal sensitivity are altered in many chronic pain conditions,¹⁰ and allodynia and spontaneous pain tend to be correlated in human studies^{11,12} and in some¹³ but not all¹⁴ mouse studies. Furthermore, sensory profiling is useful for stratifying patients in clinical trials,^{15,16} and altered sensitivity is often diagnostic.¹⁷ Ongoing pain should be assessed in addition to, not instead of, evoked pain.¹⁸ But the most problematic aspects of this testing must be rectified; for instance, outcomes of the hot water tail flick test depend more on who conducts the testing than on any other factor.¹⁹ This has received scant attention

compared with other factors, such as sex.²⁰ Outdated technology and poorly standardized testing protocols contribute to the oft-cited reproducibility crisis²¹ and are overdue for transformative improvements.

Preclinical pain tests typically measure withdrawal threshold using brief repeated (incrementing) stimuli such as von Frey filaments or sustained stimuli such as radiant heat. The stimulus intensity (force or skin temperature) at which withdrawal occurs is assumed to be the lowest intensity perceived as painful (i.e., pain threshold), notwithstanding certain caveats.²² Withdrawal might not always be triggered by pain, and focusing on threshold fails to consider pain intensity over a broader stimulus range. Recent studies have quantified responses to suprathreshold mechanical stimulation using high-speed video,^{23–25} but despite precise response measurement, stimuli were delivered by hand, and throughput was low. Resolving subtle changes in pain sensitivity requires that stimulus-response relationships be measured with high resolution (which requires reproducible stimulation and precise response measurement), over a broad dynamic range, and



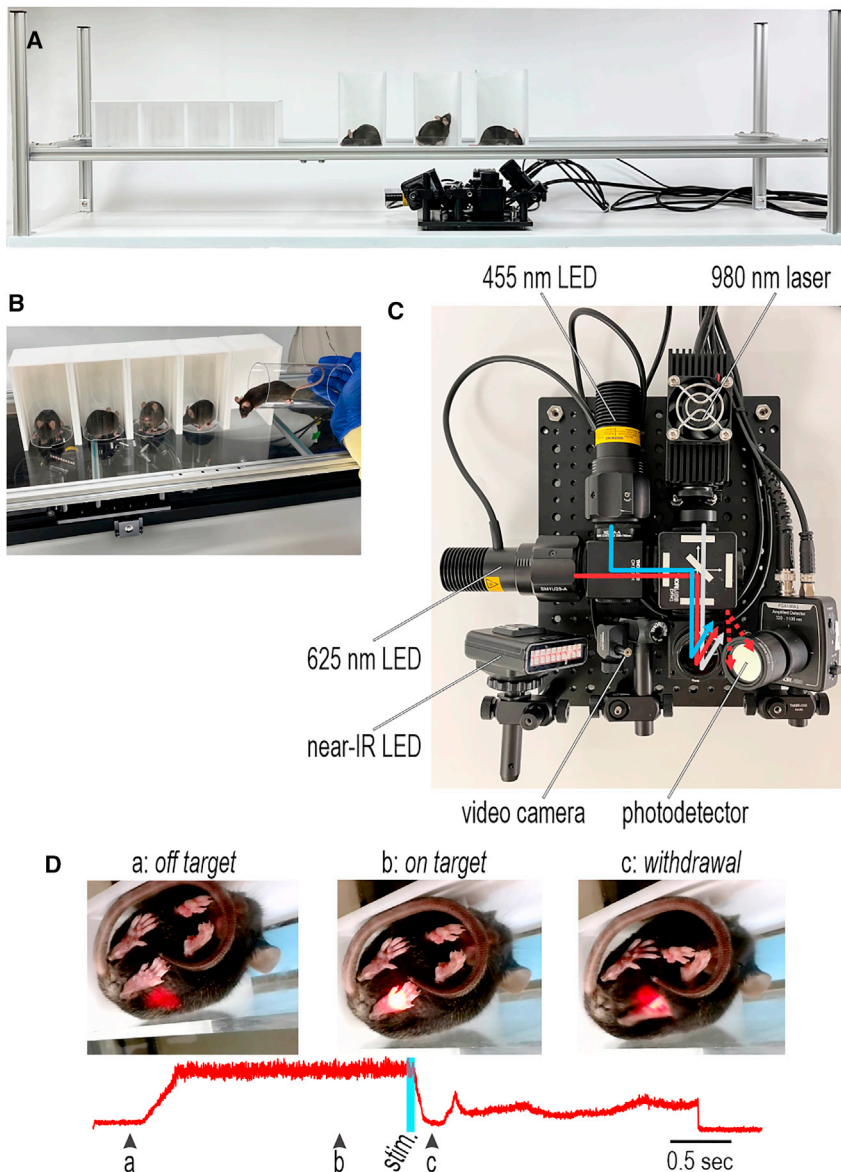


Figure 1. Overview of equipment

(A) Mice are kept separately in enclosures on a plexiglass platform (or on a metal grate for mechanical stimulation; see Figure 5B). The stimulator remains at a fixed distance below the platform and is positioned by hand (or by motorized actuators; see Figure 7).

(B) Enclosure design. Mice are transferred from their home cage to the platform in a clear plexiglass tube, which is placed vertically on the platform and slid into an opaque cubicle.

(C) Top view of photostimulator. Blue and IR light is used for optogenetic stimulation and radiant heating, respectively. Red light is used to help aim and to detect paw withdrawal. All wavelengths are combined and delivered to the same spot, but their intensities are independently controlled by computer. See Figure S1 for details.

(D) Sample frames from substage video (Video S1) before the stimulator is properly aimed (a), after aiming (b), and after paw withdrawal (c). Red trace shows the intensity of reflected red light measured by the photodetector. See also Table S1 and Video S1.

pressing Chr2 selectively in HTMRs that HTMRs can be activated in isolation.²⁸ Causal relationships between afferent co-activation patterns and perception/behavior²⁹ can be thoroughly tested in this way. Elucidating those relationships is key to understanding physiological pain and how pathology disrupts normal coding. Optogenetics has been used for basic pain research but, despite its potential, has not yet been adopted for drug testing.³⁰ Transcutaneous photostimulation is amenable to high-throughput testing but, like tactile and thermal stimuli, is hard to apply reproducibly in behaving animals.

We sought to improve preclinical pain testing by developing a device able to deliver optogenetic, thermal, and mechanical

(tactile) stimuli consistently and measure withdrawal latency automatically with millisecond precision. Using this device, we show that withdrawal latency correlates inversely with the intensity of optogenetic pulses, and that optogenetic ramps reveal differences not seen with pulses. With a clear view of the mouse from below, a neural network was trained to recognize the paw and aim the stimulator, thus fully automating the testing process. The substage video provides a wealth of data about non-reflexive behaviors for consideration alongside withdrawal measurements to more thoroughly assess the pain experience.

with reasonable efficiency (throughput). Improvements in one factor may come at the expense of other factors. The best compromise depends on the particular experiment, but improving reproducibility and throughput would be a huge benefit. Optogenetics has provided an unprecedented opportunity to study somatosensory coding, including nociception. Expressing actuators such as channelrhodopsin-2 (Chr2) in genetically defined subsets of afferents allows those afferents to be selectively activated or inhibited with light applied through the skin (transcutaneously) or directly to the nerve or spinal cord using more invasive methods.^{26,27} Afferents can be optogenetically activated in combinations not possible with somatosensory stimulation; for instance, mechanical stimuli that activate A δ high-threshold mechanoreceptors (HTMRs) normally also activate low-threshold mechanoreceptors (LTMRs), so it is only by ex-

perimental (tactile) stimuli consistently and measure withdrawal latency automatically with millisecond precision. Using this device, we show that withdrawal latency correlates inversely with the intensity of optogenetic pulses, and that optogenetic ramps reveal differences not seen with pulses. With a clear view of the mouse from below, a neural network was trained to recognize the paw and aim the stimulator, thus fully automating the testing process. The substage video provides a wealth of data about non-reflexive behaviors for consideration alongside withdrawal measurements to more thoroughly assess the pain experience.

RESULTS

Mice are kept individually in enclosures on a clear platform with the stimulator underneath (Figure 1A). A wire grate floor is used when testing mechanical stimuli (see Figure 5B). Different

enclosures were tested including a newly developed design in which the mouse is transferred from its home cage in a clear plexiglass tube, which is then turned vertically and slid into an opaque ceilinged cubicle for testing (Figure 1B). Tube/tunnel handling is less stressful than other handling methods.^{31–33} For high-speed video, which required the mouse to face left to view the stimulated paw in profile, we used a narrow rectangular chamber with clear walls on the front and left side.

Figure 1C shows the stimulator viewed from above. Blue light for optogenetic activation using ChR2, infrared (IR) light for thermal stimulation (radiant heating), and red light for aiming and response measurement are combined into a single beam using dichroic mirrors (Figure S1). The beam is directed vertically and focused to a spot 5 mm in diameter on the platform above. An adjacent camera collects video from below (substage), while a photodetector measures red light reflected off the paw (Figure 1C). A near-IR light-emitting diode (LED) helps improve lighting during high-speed video. The stimulator is translated manually or by motorized actuators (see Figure 7) using substage video to aim. For mechanical stimulation, a computer-controlled indenter is positioned below the wire grate floor (see Figure 5B) but manual or motorized/automated aiming is the same as described above.

As all wavelengths converge on the same spot, red light is turned on prior to initiating photostimulation (with blue or IR light) to verify where photostimuli will hit, thus providing visual feedback to optimize aiming (Figure 1D; Video S1). Rodents are typically assumed not to see red light³⁴; though some evidence contradicts this,^{35,36} we never observed any behavioral response to red light, suggesting that the aiming phase does not provide mice any visual cue about the forthcoming photostimulus. Reflectance of red light off the paw is measured by the adjacent photodetector (red trace). Maximization of the reflectance signal can be used to optimize aiming (compare frames a and b). This reflectance signal is stable while the paw and stimulator are immobile but changes when the paw is withdrawn (frame c), thus enabling measurement of withdrawal latency (see Figure 3). Though too slow to accurately measure fast withdrawals, standard video provides a visual record to rule out gross errors in reflectance-based latency measurements and enables assessment of slower behaviors (see Figure 6).

Reproducible stimulation

Unaccounted for variations in stimulation fundamentally limit the precision with which stimulus-response relationships can be characterized. LEDs and lasers offer stable light sources but the amount of light hitting a target can vary over time or across trials depending on the accuracy and precision of aiming. When applying light by handheld fiber optic (as typically done for transcutaneous stimulation), stability of the tester and differences in aiming technique across testers are important. To gauge the importance of aiming, we measured how the amount of light hitting a target depended on the fiber optic's positioning in the x-y plane and its distance (z) below the platform. Light was delivered through a paw-shaped cut-out to a photodiode facing downward on the platform (to simulate stimulation of a mouse paw) while controlling fiber optic position with linear actuators. Figure 2A shows that light delivery is sensitive to positioning in

all three axes, especially in z (because light rays diverge from the fiber optic tip).

To explore the practical consequences of this, we measured light delivery while 13 testers applied a 10-s-long photostimulus by handheld fiber optic (Figure 2B, blue traces). The signal-to-noise ratio ($SNR = \text{mean}^2/SD^2$) of 23.5 ± 2.0 dB (group mean \pm SEM) was significantly less than the 55.6 dB obtained with the stimulator (red trace) ($T_{12} = 15.8$, $p < 0.001$, one-sample t test). The mean stimulus intensity also differed across testers, with an inter-tester coefficient of variation ($CV = SD/\text{mean}$) of 18.8%, which is even larger than the average intra-tester CV of 8.9%. In other words, during a sustained photostimulus, temporal variations in light on target arise from each tester's instability, but this variability is compounded by differences in aiming technique across testers.

The same issues affect short (pulsed) stimuli but manifest as trial-to-trial variations. To measure variability across trials, five testers used a handheld fiber optic or the photostimulator to deliver ten 100-ms-long pulses to a photodiode (Figure 2C); each pulse was triggered independently. Trial-to-trial deviation of each tester from their individual mean dropped from 26.3 ± 3.0 mW (mean \pm SEM) with the handheld fiber optic to 6.4 ± 0.7 mW with the stimulator ($T_{98} = 6.42$, $p < 0.001$, unpaired t test), which represents a 75.6% reduction in intra-tester variance. Deviation of each tester from the group mean fell from 44.5 ± 18.2 mW with the fiber optic to 9.7 ± 2.0 mW with the photostimulator ($T_8 = 1.94$, $p = 0.093$), which represents a 78.2% reduction in inter-tester variance. In other words, using the photostimulator increased reproducibility of stimulation across testers and within each tester.

Even if stimulation is reproducible, behavior is still variable, especially in response to weak stimuli. Threshold is defined as the stimulus intensity at which withdrawal occurs on 50% of trials. Figure 2D shows determination of optogenetic threshold. Reliable aiming combined with precisely controllable LEDs (whose output can be varied in small increments over a broad range) allows one to measure threshold and characterize the broader stimulus-response relationship, assuming responses can be measured precisely.

Precise response measurement

High-speed video is the gold standard for measuring fast behaviors, but acquiring and analyzing those data are complicated and costly. We sought to replace high-speed video by detecting changes in the amount of red light reflected off the paw (see Figure 1D) using a low-cost photodetector. To validate our method, response latency was determined from high-speed video for comparison with latency determined from the reflectance signal on the same trials (Figure 3A). The stimulated paw was identified (purple dot in sample frames) using DeepLabCut³⁷ and paw height was measured from each frame (Video S2). Latency was determined independently for each signal based on the time taken for that signal to cross a threshold defined relative to the pre-stimulus baseline (see STAR Methods); after choosing a threshold value in pilot tests, the same value was applied for all subsequent measurements. Each data point in Figure 3B shows the reflectance-based and height-based latency measurement from a single trial plotted relative to one another; data are from

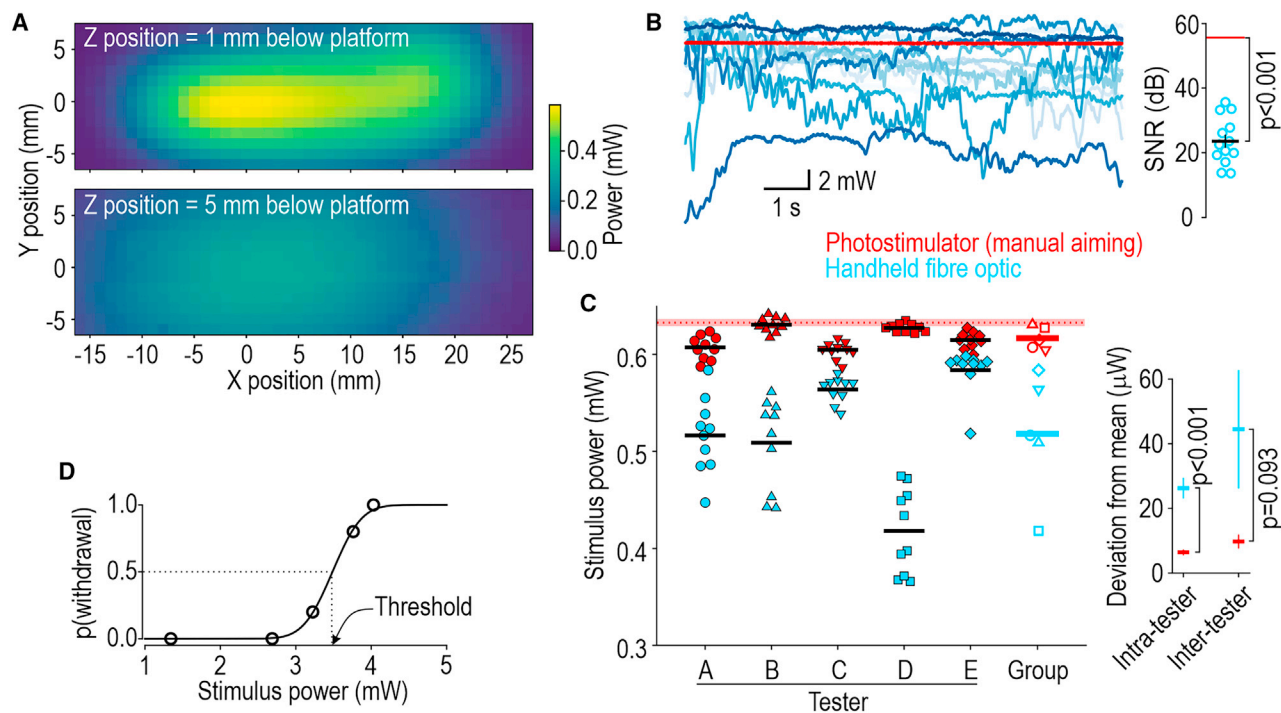


Figure 2. Reproducible photostimulation

For (A)–(C), on-target light was measured by stimulating a photodiode facing down on the platform with a paw-shaped cut-out over its surface.

(A) Importance of fiber optic positioning. Fiber optic was mounted on linear actuators to control x-y position; the x axis aligns with the long axis of the paw, with 0 position centered on the maximal response. Measurements were repeated for two distances (z) below the platform.

(B) Stability of light delivery during 10-s-long photostimuli. Signal-to-noise ratio ($SNR = \text{mean}^2/SD^2$) when using the stimulator (55.6 dB; red trace) was significantly higher than for the handheld fiber optic (23.5 ± 2.0 dB, group mean \pm SEM; blue traces, 1 for each of 13 testers) ($T_{12} = 15.8$, $p < 0.001$, one-sample t test).

(C) Trial-to-trial variability of light delivery during 100-ms-long pulses. Filled symbols show individual trials with a handheld fiber optic (blue) or stimulator (red) ($n = 5$ testers, 10 trials/tester/method); black lines represent intra-tester averages. Average trial-to-trial deviation from each tester's average was significantly larger for handheld fiber optic (26.3 ± 3.0 mW; mean \pm SEM) than for photostimulator (6.4 ± 0.7 mW) ($T_{98} = 6.42$, $p < 0.001$, unpaired t test). Open symbols represent intra-tester averages; colored lines represent group average. Tester-to-tester deviation from group average was larger for handheld fiber optic (44.5 ± 18.2 mW) than for stimulator (9.7 ± 2.0 mW) ($T_8 = 1.94$, $p = 0.093$). Red dotted line and shading show average light intensity \pm SD across 10 trials without moving the photostimulator.

(D) Example input-output curve from one mouse. Five 100-ms-long blue pulses were delivered at each of 5 intensities. Threshold is intensity at 50% probability of withdrawal, as inferred from fitted curve.

6 mice given 100-ms-long blue pulses with intensities spanning a broad range. The regression line (green, slope = 1.007) follows the equivalence line (dashed, slope = 1). Transforming these data to a Bland-Altman plot (Figure 3C) shows there is no fixed bias and that any proportional bias is inconsequential. Furthermore, the error rate is low ($<2\%$) for each method (Figure S2). Beyond avoiding an expensive high-speed camera and the challenges of filming the mouse in profile to assess paw height, the reflectance signal can be processed in real time to enable closed-loop (automated) termination of photostimuli once paw withdrawal is detected; this was implemented in most experiments using prolonged stimuli, as noted in relevant figure legends. Though not essential for latency measurements, high-speed video can provide additional information.²⁴

Characterizing stimulus-response relationships

Minimizing variability in stimulus delivery and response measurement maximizes discrimination of small biological differences; indeed, an input-output relationship is obscured by poorly

controlled input or poorly measured output adding noise respectively to the x- and y-positions of constituent data points. To explore how well our device reveals stimulus-dependent variations in withdrawal latency, we titrated the intensity of 100-ms-long pulses of blue light to determine the optogenetic threshold in each of 10 mice. Then, using intensities at defined increments above each mouse's threshold, we measured withdrawal latency as a function of photostimulus intensity (Figure 4A). Responses evoked by near-threshold intensities (blue) occurred with long latencies (>75 ms), but small increments in intensity (orange and green) evoked responses whose latencies were bimodally distributed and larger increments (red and purple) evoked short-latency (<75 ms) responses. The proportion of slow and fast responses varied significantly with photostimulus intensity ($\chi^2 = 105.0$, $p < 0.0001$, excluding purple data points). Browne et al.³⁸ reported a similar bimodal distribution of latencies but did not relate this to photostimulus intensity; instead, fast or slow responses occurred randomly in their experiments, perhaps because their ultra-short pulses (3 ms) activated

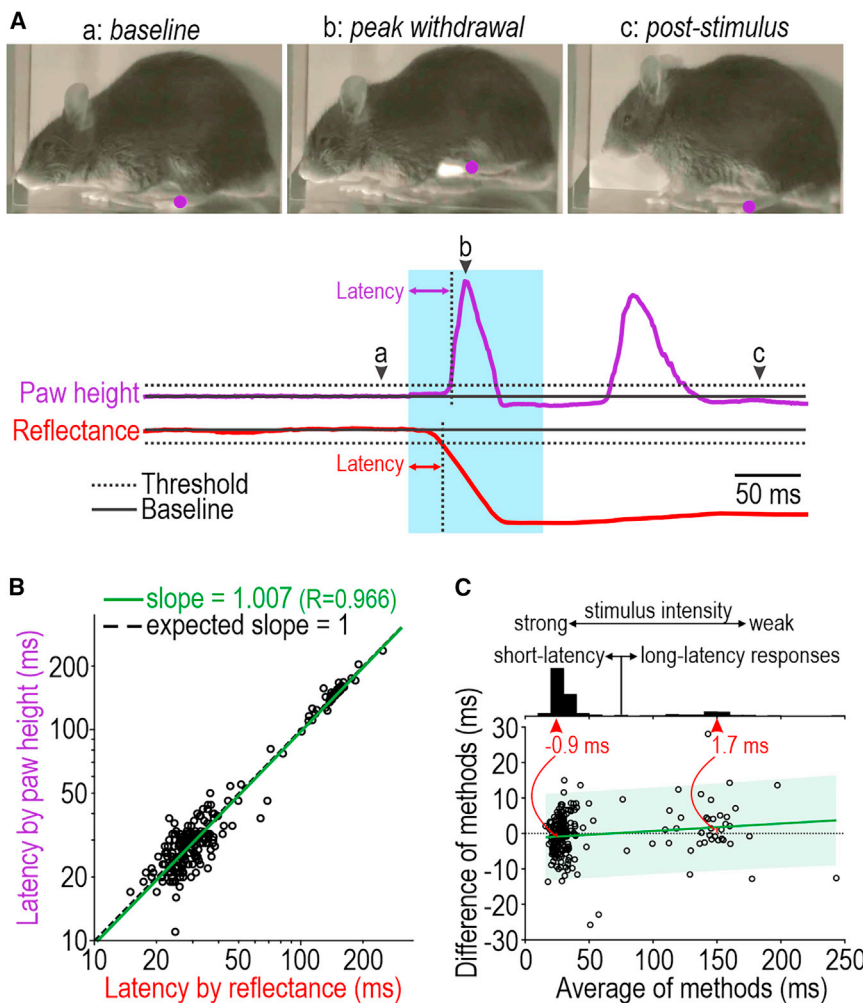


Figure 3. Precise measurement of withdrawal latency

(A) Paw withdrawal latency was measured by two methods. Paw height (purple) was extracted from high-speed video (1,000 fps) and is plotted alongside intensity of reflected red light (red) measured by the substage photodetector (1 kHz). Sample frames are shown before (a), during (b), and after (c) stimulation with position of stimulated paw (as tracked by DeepLabCut) summarized by a purple dot (Video S2). Withdrawal latency was determined as delay from stimulus onset until signal crossed threshold (dotted line) defined relative to baseline (solid line) (see STAR Methods). The same threshold value was used for all trials.

(B) Comparison of latencies measured from each signal. Data points, each representing a single trial, fell along a line representing equivalence (dashed, slope = 1), yielding a regression line (green) with slope = 1.007. Data are shown on a log scale. Data here are from 6 Advillin-ChR2 mice each tested with a range of stimulus intensities. Starting with 218 trials, 7 trials were excluded (3.2%) based on errors identified by visual inspection of raw data (see Figure S2).

(C) Bland-Altman plot using data from (B). For each trial, the difference in latency between methods (i.e., reflectance – paw height) is plotted against the average across methods. The average difference of -0.4 ms does not deviate significantly from 0 ($T_{210} = -0.8992$, $p = 0.374$, one-sample t test), meaning that there is no fixed bias. Green line and shading show regression line and 95% prediction band. Its slope (0.021) deviates significantly from horizontal ($T_{209} = 2.31$, $p = 0.021$), suggesting a proportional bias, but the errors are inconsequential: for a short-latency response of 25 ms, latency by reflectance is on average 0.9 ms shorter than latency by paw height (a 3.6% error) whereas for a long-latency response

of 150 ms, latency by reflectance is on average 1.7 ms longer (a 1.1% error) (see red highlights). Histogram shows bimodal distribution of latencies and the relation with stimulus intensity (see Figure 4).

See also Video S2.

neurons more variably despite their high intensity (47 mW/mm^2 , which is >100 times stronger than thresholds we measured using 100-ms-long pulses). Response latencies within each group decreased with increasing photostimulus intensity.

To date, all reports of withdrawal from transcutaneous optogenetic stimulation used a pulse of blue light or pulse trains,^{28,38–49} with one exception, which used sustained light to activate keratinocytes.⁵⁰ Yet different photostimulus waveforms may reveal different information about the neural control of behavior (see discussion), so testing with different waveforms will provide greater information than testing with any one waveform. Therefore, capitalizing on the stability of our stimulator (see Figure 2B), we tested slowly ramped photostimuli (and pulses) in two transgenic mouse lines: Advillin-ChR2 mice express ChR2 in all somatosensory afferents,⁵¹ whereas $\text{Na}_v1.8$ -ChR2 mice express ChR2 selectively in nociceptors.^{52,53} Whereas $\text{Na}_v1.8$ -ChR2 mice responded consistently with a latency of 1.1 ± 0.2 s (mean \pm SD), Advillin-ChR2 mice responded with much longer latencies on some trials (Figure 4B, top). The difference was due mostly to intra-mouse

variability, with individual Advillin-ChR2 mice responding with a broad range of latencies rather than some mice being consistently slow and others being consistently fast. By comparison, both genotypes exhibited a similar bimodal latency distribution when tested with pulses (Figure 4B, bottom). Latencies are nearly three orders of magnitude slower for ramp-evoked responses than for pulse-evoked responses, meaning “slow” pulse-evoked responses are still much faster than “fast” ramp-evoked responses. Our goal here was not to compare pulse and ramp stimuli but, rather, to show that one stimulus waveform might reveal differences (e.g., between genotypes) that are not revealed by other waveforms, attesting to the value of testing with different stimulus kinetics in addition to different stimulus intensities (see Figure 4A) and modalities (see below).

Additional stimulus modalities

Despite focusing hitherto on optogenetic stimuli, our device can deliver more conventional stimuli and automatically measure withdrawal. Radiant heat is applied with an IR laser (see

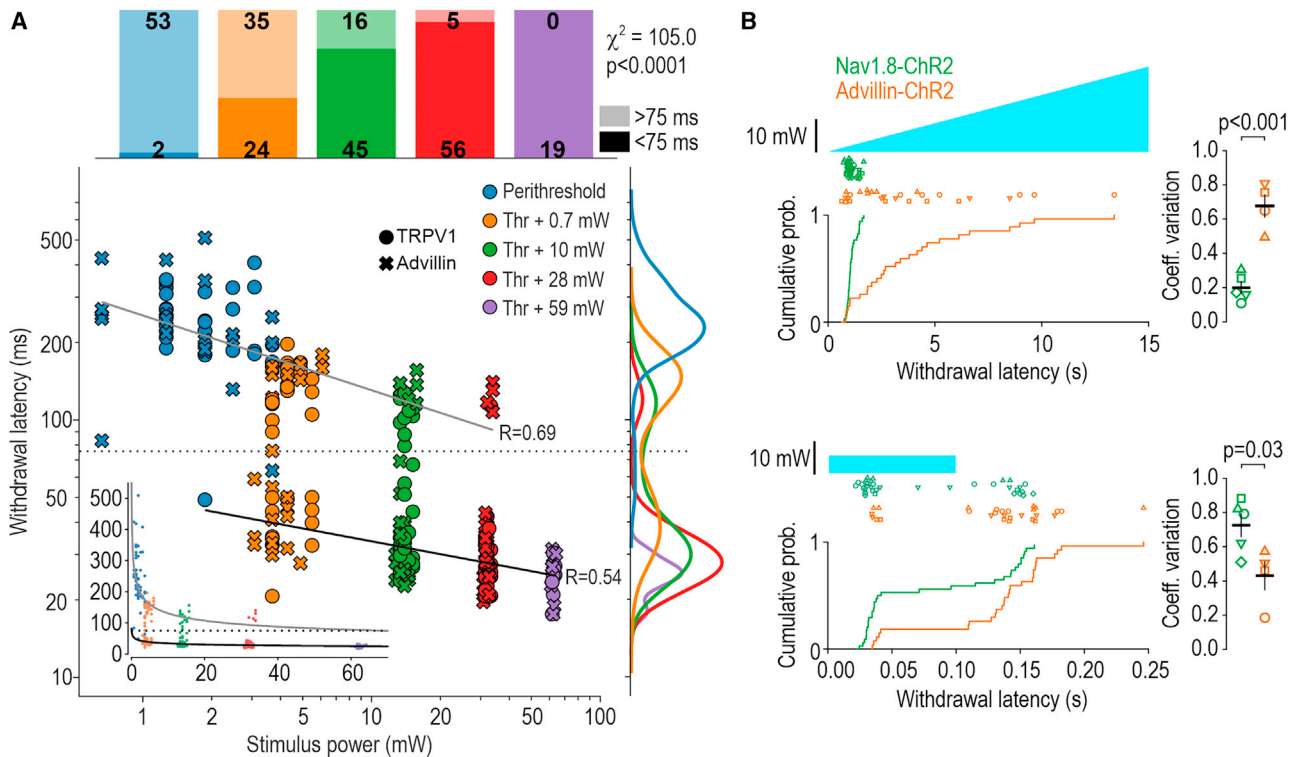


Figure 4. Stimulus-response characterization

(A) Impact of photostimulus power. Threshold intensity for 100-ms-long blue photostimuli was determined in 5 Advillin-ChR2 mice (crosses) and 5 TRPV1-ChR2 mice (circles) by two experimenters. Photostimuli were applied at increments above threshold (3 trials/mouse/suprathreshold intensity except for highest intensity, which was tested only once) and withdrawal latency was recorded. Horizontal variations in blue data points reflect intensities used for threshold determination (and are accentuated by the log scale); horizontal variations for other colors reflect inter-animal variability in threshold. Inset shows data on linear scales. Latencies exhibit a bimodal distribution (see histogram on right) with a preponderance of slow (>75 ms) responses at low intensities and fast (<75 ms) responses at high intensities (see stacked bars at top); proportions varied significantly with stimulus intensity ($\chi^2 = 105.01$, $p < 0.0001$). Black and gray lines show separate linear regressions for fast and slow responses, respectively, and correspond to exponential curves on linear scales (see inset). Fast, $y = 49.8x^{-0.17}$. Slow, $y = 254.8x^{-0.29}$.

(B) Photostimulus ramps. Nav1.8-ChR2 mice ($n = 5$, green) and Advillin-ChR2 mice ($n = 4$, orange) were tested (7 trials/mouse) with 15-s-long ramps (top; Video S3). Each mouse is represented by a different symbol. Latencies are summarized by their cumulative probability distribution. Ramp-evoked responses were significantly more variable in Advillin-ChR2 mice ($D = 0.714$, $p = 3.67 \times 10^{-8}$, two-sample Kolmogorov-Smirnov test). Variability occurs within each mouse; intra-mouse coefficient of variation (=SD/mean) was significantly higher in Advillin-ChR2 mice ($T_7 = -6.575$, $p < 0.001$, two-sample t test). Insets report mean \pm SEM. Unlike their dissimilar responses to ramps, both genotypes exhibited a similar bimodal latency distribution to pulses (bottom). Distributions differed significantly between genotypes ($D = 0.403$, $p = 0.009$) but primarily due to the different ratio of short- and long-latency responses. Intra-mouse coefficient of variation was high in both genotypes, but slightly higher in Nav1.8-ChR2 mice ($T_7 = 2.71$, $p = 0.030$). See also Video S3.

Figure 1C). Laser intensity was adjusted in pilot experiments to evoke withdrawal after ~ 8 s, as in a standard Hargreaves test. Heating was automatically terminated upon detection of paw withdrawal or after a 20 s cutoff. Withdrawal latency was significantly reduced after injecting 0.5% capsaicin into the hind paw (Figure 5A; Video S4). Mechanical stimulation is applied with a computer-controlled, force-feedback indenter (Figure 5B, inset). With the mouse positioned on metal grate floor (instead of plexiglass), the indenter tip is aimed by substage video before being raised at a fixed rate (Figure 5B; Video S5). Withdrawal is evident from the rapid drop in force as sensed by the indenter (at 1 kHz) and verified by high-speed video (at 1,000 fps), thus precluding the need for reflectance-based latency measurements. Mechanical threshold is taken as the peak force immediately prior to withdrawal.

Additional response measures

For all stimulus modalities, the substage video enables slower, non-reflexive behaviors to be analyzed. For example, video of withdrawals reported in Figure 5A revealed that thermal stimulation triggered significantly more licking, guarding, and flinching after capsaicin (Figure 6A). By plotting the occurrence or absence of these non-reflexive behaviors against latency of the preceding withdrawal, logistic regression revealed that guarding was not correlated with withdrawal latency under baseline conditions but, after capsaicin, was significantly more likely following short-latency withdrawals (Figure 6B) (logistic regression, $p = 0.605$ on the basis of 22 baseline trials vs. $p = 0.00985$ on the basis of 32 +capsaicin trials). Equivalent analysis for licking and flinching is reported in Figure S3. One may cautiously interpret this to mean capsaicin causes heat to be

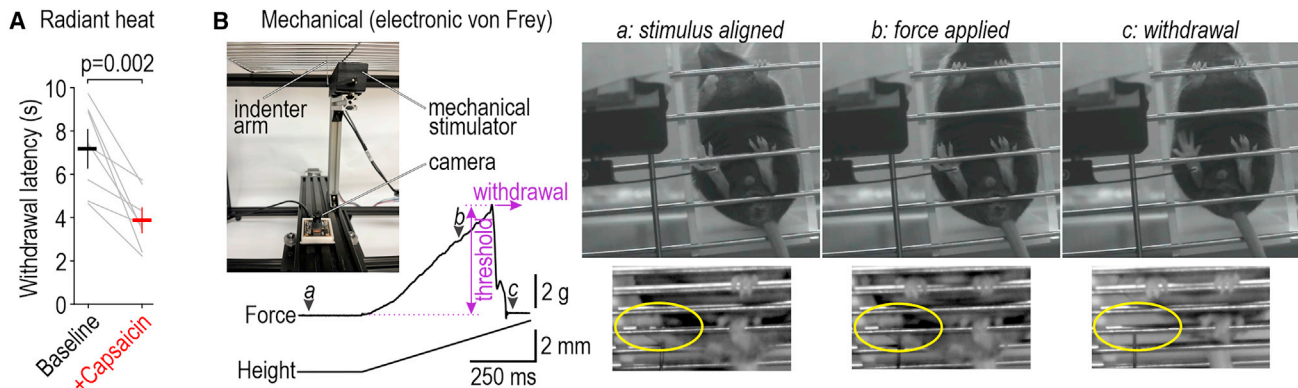


Figure 5. Additional stimulus modalities

(A) Thermal stimulation. Radiant heat was applied via infrared laser (Video S4). Latency dropped from 7.18 ± 0.74 s (mean \pm SEM) at baseline to 3.88 ± 0.49 s after injecting 0.5% capsaicin into the left hind paw ($T_7 = 4.64$, $p = 0.002$, paired t test, $n = 8$ mice, 3 trials/mouse for baseline, 3–5 trials/mouse for +capsaicin). Thermal stimulation was automatically terminated after detection of withdrawal or after 20 s cutoff.

(B) Mechanical stimulation. Top left image shows configuration of equipment. Tip of indenter arm is aimed using substage video. Height of indenter arm is ramped up by computer control while monitoring (at 1 kHz) the force exerted on the paw. Withdrawal causes an abrupt drop in force. Threshold is taken as the peak force preceding withdrawal. Sample frames extracted from the standard-rate video used for aiming (top) and the high-speed (1,000 fps) video used for validation (bottom) illustrate aiming (a), stimulation (b), and withdrawal (c) phases (Video S5). Yellow highlights the target paw, which has moved out of view in (c). See also Videos S4 and S5.

perceived as more painful, triggering faster withdrawal, whereas short-latency responses occasionally occur under baseline conditions but not because certain trials are more painful than other trials, providing clues as to where variability arises.⁵⁴

We similarly analyzed responses to optogenetic ramps because the high intra-mouse variability in withdrawal latency affords an ideal opportunity to test if short- or long-latency withdrawals in the same mouse are more or less painful. Plotting the amount of time spent licking or guarding (during a ~ 2 min post-stimulus period) against withdrawal latency (Figure 6C) revealed that long-latency withdrawals were associated with significantly more licking ($T_{55} = 5.06$, $p = 4.96 \times 10^{-6}$, one-sample t test on slope) but not more guarding ($T_{55} = -1.62$, $p = 0.111$). This suggests that failure to withdraw promptly, for reasons that remain unclear, results in the stimulus causing more pain, as inferred from licking. Interestingly, plotting time spent licking against time spent guarding on a trial-by-trial basis shows that mice tend to exhibit one or the other behavior on a given trial (Figure 6D). Interpretations warrant caution but automated analysis can expedite and help standardize future investigation along these lines.

Fully automated testing

Next, we mounted the stimulator on linear actuators (Figure 7A) so that aiming could be controlled remotely by joystick (i.e., without the tester operating in close proximity to the mice) or automatically using machine learning. For the latter, a neural network was trained with substage video using DeepLabCut to recognize the paws and other points on the mouse. The computer is then fed substage video and DeepLabCut-Live⁵⁵ uses the trained network to position the photostimulator by minimizing x- and y-error signals (to within 3 pixels) so that the target paw is positioned in the crosshairs for stimulation (Figure 7B; Video S6). Stimulation initiates

automatically once the paw has remained stationary for a minimum period and terminates automatically upon detection of paw withdrawal or after a pre-set cutoff. Automated aiming delivered stimuli even more reproducibly than manual aiming with the same device (Figure 7C). This testing was conducted using a neural network trained to recognize the paw-shaped cut-out over the photodiode; neural networks trained to recognize real paws may differ in performance, though post hoc analysis of video records suggests excellent performance identifying the center of the paw and targeting that point for stimulation. For mechanical stimulation, the metal grate floor partially obscures the mouse but automated aiming is still possible with an appropriately trained neural network (Video S7).

After completing a trial, the device automatically proceeds to the neighboring mouse. By interleaving trials, other mice are tested during the inter-stimulus interval required for each mouse, thus expediting the overall testing process. The order of testing can easily be randomized, which is difficult for an experimenter to keep track of. Meta-data (mouse identification, date, time, stimulus parameters), raw data (video, reflectance signal) and measurements (latency, threshold) are automatically saved (Figure 7E; Data S1). Non-reflexive behaviors such as those analyzed in Figure 6 can also be detected and quantified automatically using unsupervised methods (Figure 7F; Video S8). Although not yet implemented, automated real-time classification of posture can be integrated into withdrawal testing so that stimuli are applied contingent on certain postures; for example, guarding and rearing can influence withdrawal latencies^{38,56} and stimulation could, therefore, be delayed until the mouse assumes a preferred posture. Even without closed-loop control, video records enable post hoc correlation of withdrawal latency with the pre-stimulus posture. Standardized high-throughput testing without the potential errors, systematic differences,¹⁹

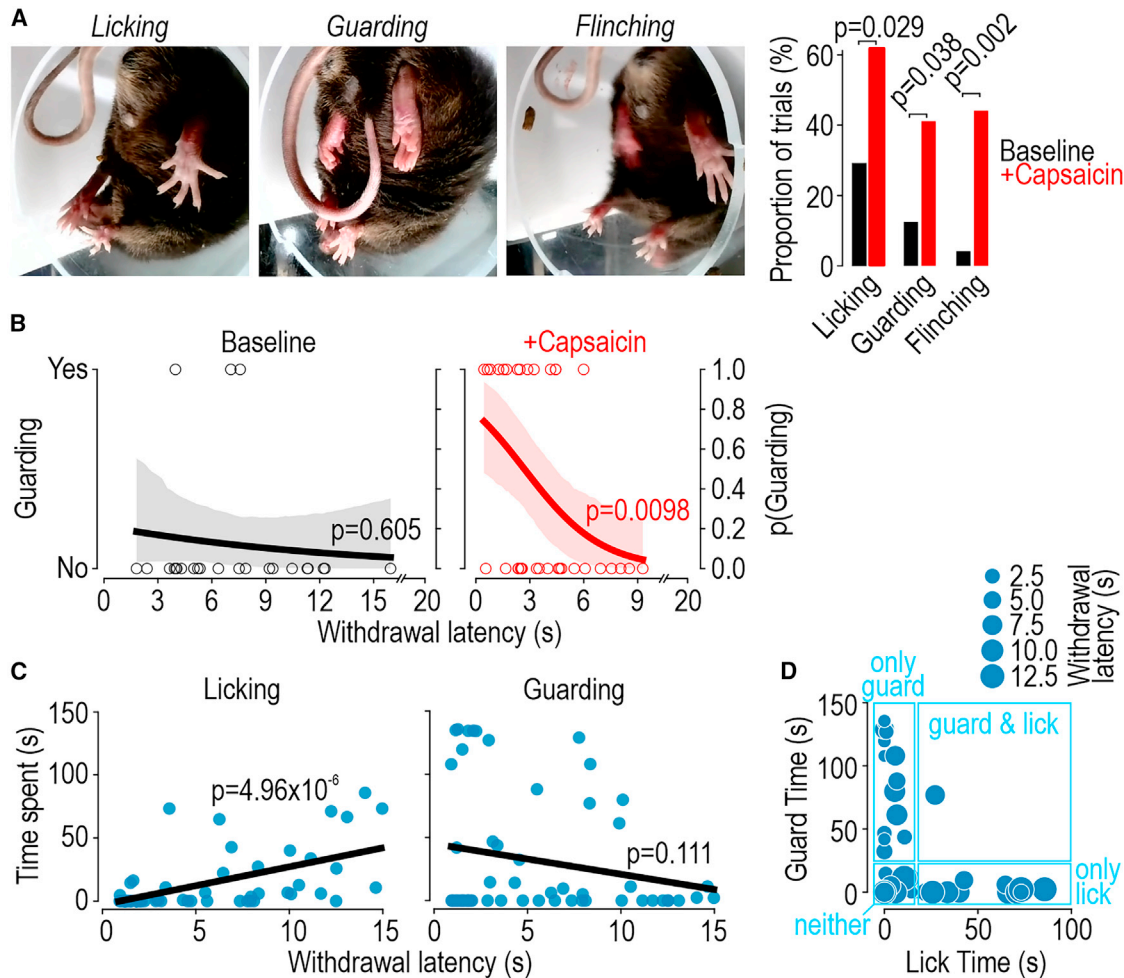


Figure 6. Additional behavior measurements

(A) Non-reflexive behaviors after radiant heat. Sample frames from Video S4 illustrate licking, guarding, and flinching (same experiments reported in Figure 5A). These behaviors were significantly more common after capsaicin (red; $n = 34$ trials) than at baseline (black; $n = 24$ trials). p values on graph report χ^2 tests. These behaviors were rare in the pre-stimulus period.

(B) Correlation between guarding and withdrawal latency. Extending analysis in (A), occurrence of guarding on a given trial (yes/no) was plotted against withdrawal latency. According to logistic regression, guarding was significantly more likely after shorter latency responses in the +capsaicin condition ($p = 0.0098$) but not at baseline ($p = 0.605$). Shading indicates bootstrapped 95% confidence interval. See Figure S3 for analysis of licking and flinching.

(C) Non-reflexive behaviors after optogenetic ramps. In 5 TRPV1-ChR2 mice given 15-s-long optogenetic ramps (9–12 trials/mouse), withdrawal latency was plotted against the time spent licking or guarding during the subsequent post-withdrawal period (2 min duration + time remaining in ramp after withdrawal). p values on graphs show strength of correlation, which trended in opposite directions. Optogenetic stimulation was automatically terminated after detection of withdrawal.

(D) Guarding vs. licking. Plotting guard time against lick time on a trial-by-trial basis reveals that mice typically exhibit one or the other behavior on a given trial, not both.

See also Video S4.

and animal stress⁵⁷ associated with human testers is thus realized.

DISCUSSION

We developed a device able to reproducibly deliver photostimuli of different wavelengths, intensities, and kinetics (waveforms). A photometer detects paw withdrawal and measures withdrawal latency with millisecond precision based on changes in the reflectance of red light. The accuracy of this approach was vali-

dated by comparison with high-speed video. Closed-loop control of stimulation is made possible by real-time detection of paw movement. We also demonstrate computer-controlled mechanical stimulation and automated detection of touch-evoked withdrawal. Building on computer-controlled stimulation and response measurement, we automated video-based aiming by using neural networks to track the paw plus motorized actuators to move the stimulator. Whereas aiming by joystick prevents the tester from working in close proximity to the mice, which stresses them,⁵⁷ automation removes the human element

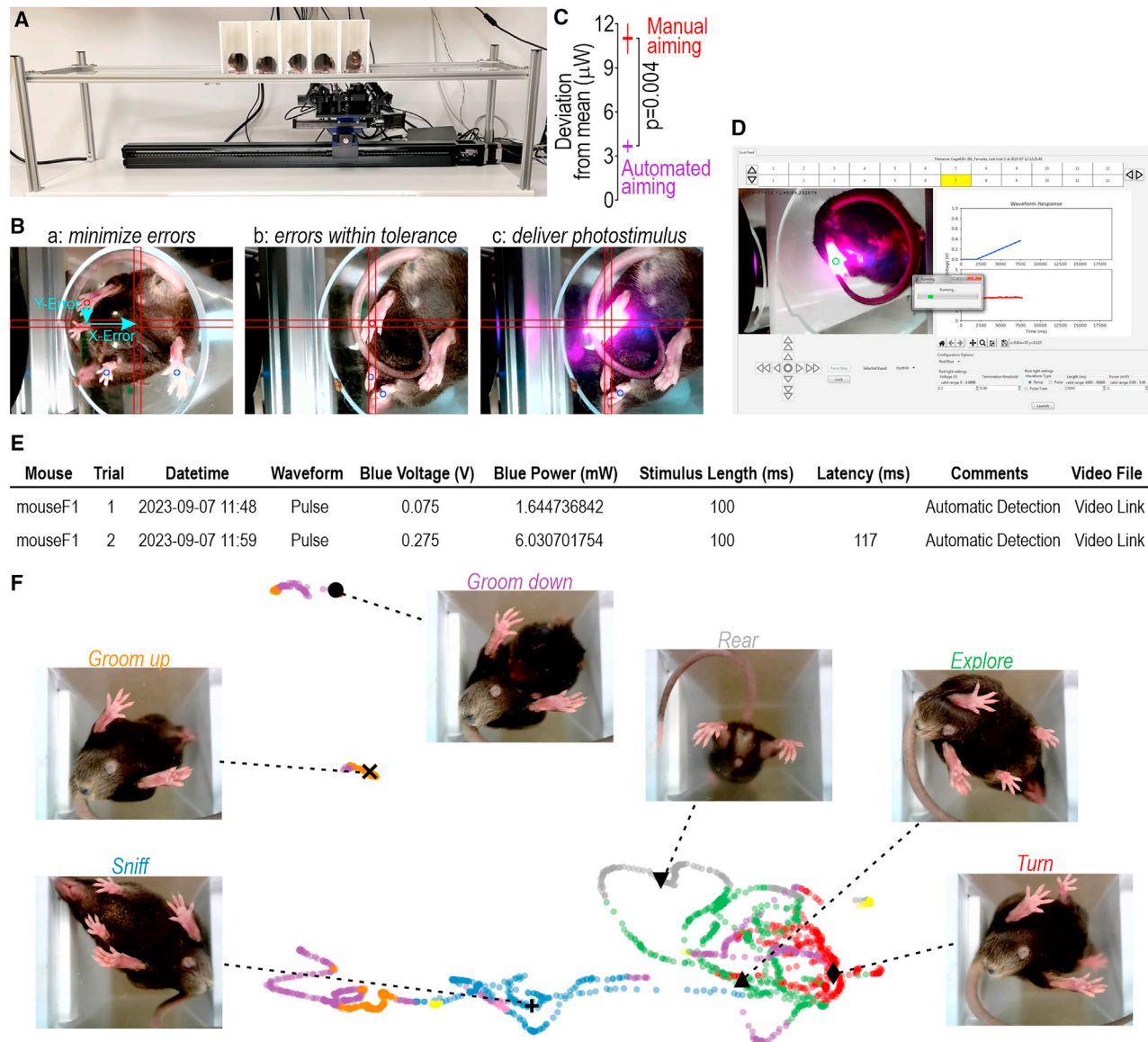


Figure 7. Fully automated testing

(A) Motorized photostimulator. Unlike the manually aimed version (see Figure 1), this version is mounted on linear actuators that allow its computer-controlled translation in the x and y axes. Aiming is controlled by joystick or automatically with machine learning.

(B) Automated aiming. A neural network identifies the paws, snout, and tail base from substage video. Deviation of the target paw from the crosshairs is measured (a) and then minimized by translating the stimulator (b). In other words, x- and y-errors are reduced until the center of the target paw (red dot) is within 3 pixels of the screen center, which is aligned with the stimulation zone. Once the paw has remained stable inside the crosshairs for a minimum (user-defined) period, photostimulation is initiated (c), and withdrawal latency is measured and recorded. The stimulator then moves to the adjacent mouse. See Video S6 for an example of fully automated testing with radiant heat. Automated aiming is also compatible with the wire grate floor used for mechanical stimulation (Video S7).

(C) Automated aiming (10 trials), quantified as in Figure 2C, was significantly less variable than manual aiming (5 testers, 10 trials each) of the photostimulator ($T_{58} = 3.03$, $p = 0.004$, unpaired t test). Stability of prolonged photostimulation (cf. Figure 2B) is equivalent regardless of aiming method. Graph reports mean \pm SEM.

(D) Graphical user interface. Figure S4 shows enlarged view with description.

(E) Sample spreadsheet to all which data and metadata from each trial is automatically saved with linked graphs and videos. See Data S1 for a sample spreadsheet with data and linked videos for fully automated testing of optogenetic pulses and ramps.

(F) Automated, non-supervised classification of ongoing behavior. Colored dots show latent space embedding of behavioral state determined at regular intervals from Video S8. Screenshots corresponding to labeled points illustrate specific behaviors.

See also Data S1 and Videos S6, S7, and S8.

altogether, with significant benefits for standardization, objectivity, and throughput. Video records allow non-reflexive behaviors to be quantified and correlated with withdrawal measurements. Automation also facilitates standardized recording of data and metadata, which is crucial for creating large datasets amenable to mining.

Capitalizing on identification of gene expression patterns to distinguish subtypes of somatosensory afferents,⁵⁸ optogenetics affords an unprecedented opportunity to study somatosensory coding by activating or inhibiting specific afferent subtypes. Testing the behavioral response to synthetic activation patterns allows one to explore causal relationships, complementing efforts to characterize co-activation patterns evoked by natural stimuli.⁵⁹ This does not require that optogenetic stimuli mimic natural, somatosensory stimuli. Optogenetic stimuli are patently unnatural—and the evoked sensations probably feel unnatural, like paresthesias evoked by electrical stimulation—but their ability to evoke behavior allows one to start inferring how they are perceived, and how those sensations relate to neural activation patterns. Doing this requires tight control of the stimulus and precise measurement of the response.

Technical advances have been made in delivering photostimuli to the CNS or peripheral nerves for optogenetic manipulations.⁶⁰ This technology is invaluable, but stimulating the nerve does not reproduce somatotopically organized stimulation. Photostimulating receptive fields in the skin is preferable in that regard, and is also less invasive. However, transcutaneous stimuli are difficult to apply reproducibly to behaving animals. Sharif et al.⁴⁸ solved this by mounting the fiber optic to the head to stimulate the cheek, but a comparable solution is infeasible for stimulating paws. These technical challenges explain why past studies focused on whether mice responded to optogenetic stimulation, without carefully varying stimulus parameters or measuring subtler aspects of the response. Past studies have varied the number, rate, or intensity of pulses, but in the supra-threshold regime, with consequences for the amount of licking, jumping, or vocalization. To our knowledge, only one study⁶¹ titrated the intensity of transcutaneous photostimuli to determine threshold (see Figure 2D), and another⁴³ titrated pulse duration and spot size. Moreover, scoring responses by eye, though still the norm for many tests, must be replaced with objective metrics. Schorscher-Petcu et al.⁴³ recently described a device that uses galvanometric mirrors to direct photostimuli and high-speed substage video to measure withdrawal. Their device is very elegant, but reliance on high-speed video to detect withdrawals likely precludes closed-loop control, and nor is their device fully automated or high-throughput. Our device delivers reproducible photostimuli and automatically measures withdrawals using the red-reflectance signal complemented by regular-speed substage video, which is also used for automated aiming.

Non-painful stimuli may trigger withdrawal, which is to say that the threshold stimulus (or the probability of withdrawal) may not reflect painfulness.²⁴ In that respect, testing with stronger stimuli is informative. The study by Browne et al.³⁸ stands out for its use of high-speed video to thoroughly quantify responses to optogenetic stimulation. Like us (see Fig-

ure 4A), they observed a bimodal distribution of withdrawal latencies; however, they observed this despite using high-intensity pulses, most likely because their pulses were extremely brief (3 ms) and might, therefore, have activated afferents probabilistically. By varying the intensity of longer (100 ms) pulses, we observed that stronger stimuli evoke faster withdrawals, evident as a continuous shift in latency as well as a switching from long- to short-latency responses. A putative explanation for the bimodal latency distribution—consistent with Browne et al.³⁸ and with the double-alarm system proposed by Plaghki et al.⁶² based on different rates of heating—is that slow and fast responses are mediated by C- and A-fibers, respectively. Building from that, our data suggest that C-fibers are recruited first (i.e., by weaker photostimuli) and that slow responses speed up as more C-fibers get recruited, but a discontinuous “switch” to fast responses occurs once A-fibers get recruited, and fast responses speed up as more A-fibers are recruited. Further investigation is required but resolving the stimulus-response relationship sufficiently to even pose such questions is notable.

Browne et al.³⁸ also noted that the withdrawal response was not limited to the stimulated limb and, instead, was more widespread. Although not quantified here, a widespread response was evident in substage video and sometimes included vocalizing, facial grimacing, jumping, and orienting to the stimulus followed by licking, guarding, or flinching of the stimulated paw (see Figure 6). A complete analysis of each trial ought to consider not only the reflexive component (i.e., did withdrawal occur and how quickly), but also whether signs of discomfort were exhibited afterward and for how long. Those signs are obvious when applying strong stimuli but become harder to discern with near-threshold stimuli, which makes objective quantification all the more important. Regular-speed video is sufficient to capture all but the fast reflex (which can be measured by other means; see Figure 3) and the bottom-up view is well suited for machine-learning-based quantification of ongoing behaviors.^{63,64} We recommend that video be recorded for all trials if only to allow analysis of those data in the future. There has been an explosion of artificial intelligence (AI)-based methods for extracting key points on animals^{37,65,66} and algorithms for extracting higher-level behaviors from key point^{63,64,67} or raw pixel⁶⁸ data. Application of such tools is yielding impressive results.⁶⁹ Other hardware has been recently developed to facilitate such analysis but does not include stimulation capabilities.⁷⁰ By capturing video before and after stimuli, our device enables users to quantify behaviors in addition to measuring reflexive withdrawal using traditional metrics.

Withdrawal responses are known to be sensitive to posture and ongoing behavior at the time of stimulation.^{38,56,71} Such differences may confound latency measurements but may also provide important information; either way, they should be accounted for. Waiting for each mouse to adopt a specific posture is onerous for a human tester but can be done by computer if automated stimulation is made contingent on the mouse being in a certain state. Before that, to better understand the relationship between the mouse’s pre-stimulus state and its subsequent stimulus-evoked withdrawal (and post-stimulus state), its state at stimulus onset could be classified from video (like in

Figure 7F) and correlated with the evoked response on that trial (as for post-stimulus behaviors in Figures 6B and 6C). Other comparisons would also be informative, like correlating the pre- and post-stimulus states and treatment status. In short, more data can be acquired and more thoroughly analyzed than is typically done in current protocols; others have also advocated for this.^{23–25} More comprehensive analysis need not entail expensive equipment or reduced throughput.

Nearly all past studies involving transcutaneous optogenetic stimulation used single pulses or pulse trains.^{28,38–49} In the one exception, Baumbauer et al.⁵⁰ activated ChR2-expressing keratinocytes with sustained light. Single pulses and pulse trains are just some of the many possible waveforms, especially as LEDs can be so easily controlled. Notably, just like pulsed electrical stimuli lost favor in pain testing because of the unnaturally synchronized neural activation they evoke,¹ pulsed optogenetic stimuli warrant similar scrutiny and should not be the only waveform tested. Indeed, different rates of radiant heating differentially engage C- and A-fibers,⁶² thus enabling the role of different afferents to be studied. By testing photostimulus ramps, we uncovered genotypic differences that were not evident with pulses (see Figure 4B). The basis for the genotypic difference requires further investigation but we hypothesize that co-activation of non-nociceptive afferents in Advillin- and TRPV1-ChR2 mice (but not in Nav1.8-ChR2 mice) engages a gate control mechanism that tempers effects of nociceptive input, consistent with Arcourt et al.,²⁸ who showed that activating A δ -HTMRs in isolation evoked more guarding, jumping, and vocalization than co-activating A δ -HTMRs and LTMRs. By testing different photostimulus waveforms, one can start to delineate the underlying interactions. Photostimulus kinetics influence how optogenetic actuators such as ChR2 respond (e.g., whether they desensitize, thus producing less current for a given photostimulus intensity), but one must also consider how neurons respond to those photocurrents. Specifically, pulsed stimuli tend to evoke precisely timed spikes,³⁸ leading to spikes that are synchronized across co-activated neurons,⁷² which may or may not accurately reflect the spiking patterns evoked by somatosensory stimuli. Artificial stimuli need not mimic natural stimuli to be informative; indeed, deliberately evoking spiking patterns not possible with natural stimuli offers new opportunities to probe somatosensory coding, including the role of synchrony. In that respect, optogenetic testing should include but not be limited to pulsed photostimuli. Interestingly, some studies^{44,45} have tested if ChR2-expressing mice avoid blue-lit floors, which they do. In these cases, the floor light was continuous, unlike the pulses typically applied by fiber optic; it is therefore notable that mice avoided the blue floor but did not respond to it with reflexive withdrawal, paw licking, or other outward signs of pain, as they did to pulses. In the one case where the floor light was pulsed,⁴⁷ reflexive withdrawal was observed. These results highlight the underappreciated importance of stimulus kinetics.

To summarize, we describe a device capable of reproducible, automated, multimodal algometry, or RAMalgo. Aiming, stimulation, and measurement are fully automated, which improves standardization and increases throughput, among other benefits. A video record of the animal before, during, and after stimu-

lation allows one to extend analysis beyond traditional response metrics (i.e., threshold and latency) to consider if evoked and ongoing pain behaviors are correlated.

Limitations of the study

Though we have demonstrated computer-controlled mechanical stimulation, we did not compare this against von Frey filaments, for instance, by comparing intra- and inter-tester variability in stimulus delivery like we did for photostimulation. There is still room to improve automated mechanical stimulation and to add other stimulus modalities. Many other conceivable scenarios have not yet been tested; for instance, *in vivo* electrophysiology and calcium imaging could be precisely synchronized with stimulation and behavioral response measurements using this technology. Unlike home-cage monitoring, mice must be transferred to an unfamiliar environment for testing. Best practices must be applied when handling mice, acclimating them, testing at the same time of day, etc., to minimize stress and the variability it introduces.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - Photostimulator
 - Mechanostimulator
 - Platform and enclosures
 - Comparison with handheld fiber optic
 - Automated withdrawal detection and latency measurement
 - High-speed video
 - Pose estimation
 - Automated aiming
 - Behavior extraction
 - Capsaicin and heat hypersensitivity
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.crmeth.2023.100650>.

ACKNOWLEDGMENTS

This study was funded by a Foundation Grant from the Canadian Institutes of Health Research (FDN167276) and a Proof-of-Principle grant from The Hospital for Sick Children to S.A.P. and by an Ontario Graduate Scholarship and Restrcomp Award to C.D. We thank Jason Jeong for expert technical assistance, Kate Zhang for help with experiments, Erica Austriaco for contributions to the software, and volunteer testers whose participation allowed us to compare light delivery methods.

AUTHOR CONTRIBUTIONS

Conceptualization, C.D. and S.A.P.; methodology, C.D. and S.A.P.; software, C.D. and M.A.A.; hardware, C.D., M.A.A., and S.A.P.; investigation, C.D. and M.A.A.; writing – original draft, C.D.; writing – review & editing, C.D. and S.A.P.; supervision, S.A.P.; funding acquisition, S.A.P.

DECLARATION OF INTERESTS

The authors have filed a US patent (18/371.847) for the device described in this work.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: May 4, 2023

Revised: September 11, 2023

Accepted: October 30, 2023

Published: November 21, 2023

REFERENCES

- Le Bars, D., Gozariu, M., and Cadden, S.W. (2001). Animal models of nociception. *Pharmacol. Rev.* *53*, 597–652.
- Deuis, J.R., Dvorakova, L.S., and Vetter, I. (2017). Methods used to evaluate pain behaviors in rodents. *Front. Mol. Neurosci.* *10*, 284. <https://doi.org/10.3389/fnmol.2017.00284>.
- Barrot, M. (2012). Tests and models of nociception and pain in rodents. *Neuroscience* *211*, 39–50. <https://doi.org/10.1016/J.NEUROSCIENCE.2011.12.041>.
- Gregory, N.S., Harris, A.L., Robinson, C.R., Dougherty, P.M., Fuchs, P.N., and Sluka, K.A. (2013). An overview of animal models of pain: disease models and outcome measures. *J. Pain* *14*, 1255–1269. <https://doi.org/10.1016/J.JPAIN.2013.06.008>.
- Burma, N.E., Leduc-Pessah, H., Fan, C.Y., and Trang, T. (2017). Animal models of chronic pain: Advances and challenges for clinical translation. *J. Neurosci. Res.* *95*, 1242–1256. <https://doi.org/10.1002/JNR.23768>.
- Jaggi, A.S., Jain, V., and Singh, N. (2011). Animal models of neuropathic pain. *Fundam. Clin. Pharmacol.* *25*, 1–28. <https://doi.org/10.1111/J.1472-8206.2009.00801.X>.
- Abboud, C., Duveau, A., Bouali-Benazzouz, R., Massé, K., Mattar, J., Brochoire, L., Fossat, P., Boué-Grabot, E., Hleihel, W., and Landry, M. (2021). Animal models of pain: Diversity and benefits. *J. Neurosci. Methods* *348*, 108997. <https://doi.org/10.1016/J.JNEUMETH.2020.108997>.
- Mogil, J.S., and Crager, S.E. (2004). What should we be measuring in behavioral studies of chronic pain in animals? *Pain* *112*, 12–15. <https://doi.org/10.1016/J.PAIN.2004.09.028>.
- Backonja, M.M., and Stacey, B. (2004). Neuropathic pain symptoms relative to overall pain rating. *J. Pain* *5*, 491–497. <https://doi.org/10.1016/J.JPAIN.2004.09.001>.
- Maier, C., Baron, R., Tölle, T.R., Binder, A., Birbaumer, N., Birklein, F., Gierthmühlen, J., Flor, H., Geber, C., Hüge, V., et al. (2010). Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): Somatosensory abnormalities in 1236 patients with different neuropathic pain syndromes. *Pain* *150*, 439–450. <https://doi.org/10.1016/J.PAIN.2010.05.002>.
- Koltzenburg, M., Torebjörk, H.E., and Wahren, L.K. (1994). Nociceptor modulated central sensitization causes mechanical hyperalgesia in acute chemogenic and chronic neuropathic pain. *Brain* *117*, 579–591. <https://doi.org/10.1093/BRAIN/117.3.579>.
- Rowbotham, M.C., and Fields, H.L. (1996). The relationship of pain, allodynia and thermal sensation in post-herpetic neuralgia. *Brain* *119*, 347–354. <https://doi.org/10.1093/BRAIN/119.2.347>.
- Pitzer, C., Kuner, R., and Tappe-Theodor, A. (2016). Voluntary and evoked behavioral correlates in neuropathic pain states under different social housing conditions. *Mol. Pain* *12*, 1744806916656635. <https://doi.org/10.1177/1744806916656635>.
- Mogil, J.S., Graham, A.C., Ritchie, J., Hughes, S.F., Austin, J.S., Schorscher-Petcu, A., Langford, D.J., and Bennett, G.J. (2010). Hypolocotion, asymmetrically directed behaviors (licking, lifting, flinching, and shaking) and dynamic weight bearing (gait) changes are not measures of neuropathic pain in mice. *Mol. Pain* *6*, 34. <https://doi.org/10.1186/1744-8069-6-34>.
- Edwards, R.R., Dworkin, R.H., Turk, D.C., Angst, M.S., Dionne, R., Freeman, R., Hansson, P., Haroutounian, S., Arendt-Nielsen, L., Attal, N., et al. (2016). Patient phenotyping in clinical trials of chronic pain treatments: IMMPACT recommendations. *Pain* *157*, 1851–1871. <https://doi.org/10.1097/J.PAIN.0000000000000602>.
- Baron, R., Dickenson, A.H., Calvo, M., Dib-Hajj, S.D., and Bennett, D.L. (2023). Maximizing treatment efficacy through patient stratification in neuropathic pain trials. *Nat. Rev. Neurol.* *19*, 53–64. <https://doi.org/10.1038/S41582-022-00741-7>.
- Arnold, L.M., Bennett, R.M., Crofford, L.J., Dean, L.E., Clauw, D.J., Goldenberg, D.L., Fitzcharles, M.A., Paiva, E.S., Staud, R., Sarzi-Puttini, P., et al. (2019). AAPT diagnostic criteria for fibromyalgia. *J. Pain* *20*, 611–628. <https://doi.org/10.1016/J.JPAIN.2018.10.008>.
- Negus, S.S. (2019). Core outcome measures in preclinical assessment of candidate analgesics. *Pharmacol. Rev.* *71*, 225–266. <https://doi.org/10.1124/PR.118.017210>.
- Chesler, E.J., Wilson, S.G., Lariviere, W.R., Rodriguez-Zas, S.L., and Mogil, J.S. (2002). Influences of laboratory environment on behavior. *Nat. Neurosci.* *5*, 1101–1102. <https://doi.org/10.1038/NN1102-1101>.
- Sadler, K.E., Mogil, J.S., and Stucky, C.L. (2022). Innovations and advances in modelling and measuring pain in animals. *Nat. Rev. Neurosci.* *23*, 70–85. <https://doi.org/10.1038/S41583-021-00536-7>.
- Mogil, J.S. (2017). Laboratory environmental factors and pain behavior: the relevance of unknown unknowns to reproducibility and translation. *Lab Anim. (NY)* *46*, 136–141. <https://doi.org/10.1038/labani.1223>.
- Le Bars, D., Hansson, P.T., and Plaghki, L. (2009). Current animal test and models of pain. In *Pharmacology of Pain*, D. Beaulieu, F. Lussier, and A.H. Porreca, eds. (IASP Press), pp. 475–504.
- Abdus-Saboor, I., Fried, N.T., Lay, M., Burdge, J., Swanson, K., Fischer, R., Jones, J., Dong, P., Cai, W., Guo, X., et al. (2019). Development of a mouse pain scale using sub-second behavioral mapping and statistical modeling. *Cell Rep.* *28*, 1623–1634.e4. <https://doi.org/10.1016/j.celrep.2019.07.017>.
- Jones, J.M., Foster, W., Twomey, C.R., Burdge, J., Ahmed, O.M., Pereira, T.D., Wojcik, J.A., Corder, G., Plotkin, J.B., and Abdus-Saboor, I. (2020). A machine-vision approach for automated pain measurement at millisecond timescales. *Elife* *9*, e57258. <https://doi.org/10.7554/eLife.57258>.
- Fried, N.T., Chamesian, A., Zylka, M.J., and Abdus-Saboor, I. (2020). Improving pain assessment in mice and rats with advanced videography and computational approaches. *Pain* *161*, 1420–1424. <https://doi.org/10.1097/j.pain.0000000000001843>.
- Copits, B.A., Pullen, M.Y., and Gereau, R.W. (2016). Spotlight on pain: Optogenetic approaches for interrogating somatosensory circuits. *Pain* *157*, 2424–2433. <https://doi.org/10.1097/J.PAIN.0000000000000620>.
- Xie, Y.F., Wang, J., and Bonin, R.P. (2018). Optogenetic exploration and modulation of pain processing. *Exp. Neurol.* *306*, 117–121. <https://doi.org/10.1016/J.EXPNEUROL.2018.05.003>.
- Arcourt, A., Gorham, L., Dhandapani, R., Prato, V., Taberner, F.J., Wende, H., Gangadharan, V., Birchmeier, C., Heppenstall, P.A., and Lechner, S.G. (2017). Touch receptor-derived sensory information alleviates acute pain signaling and fine-tunes nociceptive reflex coordination. *Neuron* *93*, 179–193. <https://doi.org/10.1016/j.neuron.2016.11.027>.

29. Prescott, S.A., and Ratté, S. (2012). Pain processing by spinal microcircuits: afferent combinatorics. *Curr. Opin. Neurobiol.* 22, 631–639. <https://doi.org/10.1016/j.conb.2012.02.010>.
30. Woolf, C.J. (2020). Capturing novel non-opioid pain targets. *Biol. Psychiatr.* 87, 74–81. <https://doi.org/10.1016/j.biopsych.2019.06.017>.
31. Hurst, J.L., and West, R.S. (2010). Taming anxiety in laboratory mice. *Nat. Methods* 7, 825–826. <https://doi.org/10.1038/nmeth.1500>.
32. Gouveia, K., and Hurst, J.L. (2013). Reducing mouse anxiety during handling: Effect of experience with handling tunnels. *PLoS One* 8, e66401. <https://doi.org/10.1371/JOURNAL.PONE.0066401>.
33. Gouveia, K., and Hurst, J.L. (2019). Improving the practicality of using non-aversive handling methods to reduce background stress and anxiety in laboratory mice. *Sci. Rep.* 9, 20305. <https://doi.org/10.1038/s41598-019-56860-7>.
34. Rocha, F.A.d.F., Gomes, B.D., Silveira, L.C.d.L., Martins, S.L., Aguiar, R.G., De Souza, J.M., and Ventura, D.F. (2016). Spectral sensitivity measured with electroretinogram using a constant response method. *PLoS One* 11, e0147318. <https://doi.org/10.1371/JOURNAL.PONE.0147318>.
35. Nikbakht, N., and Diamond, M.E. (2021). Conserved visual capacity of rats under red light. *Elife* 10, e66429. <https://doi.org/10.7554/ELIFE.66429>.
36. Niklaus, S., Albertini, S., Schnitzer, T.K., and Denk, N. (2020). Challenging a myth and misconception: red-light vision in rats. *Animals* 10, 422. <https://doi.org/10.3390/ANI10030422>.
37. Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., and Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* 21, 1281–1289. <https://doi.org/10.1038/s41593-018-0209-y>.
38. Browne, L.E., Latremoliere, A., Lehnert, B.P., Grantham, A., Ward, C., Alexandre, C., Costigan, M., Michoud, F., Roberson, D.P., Ginty, D.D., and Woolf, C.J. (2017). Time-resolved fast mammalian behavior reveals the complexity of protective pain responses. *Cell Rep.* 20, 89–98. <https://doi.org/10.1016/j.celrep.2017.06.024>.
39. Dhandapani, R., Arokiaraj, C.M., Taberner, F.J., Pacifico, P., Raja, S., Nocchi, L., Portulano, C., Franciosa, F., Maffei, M., Hussain, A.F., et al. (2018). Control of mechanical pain hypersensitivity in mice through ligand-targeted photoablation of TrkB-positive sensory neurons. *Nat. Commun.* 9, 1640–1714. <https://doi.org/10.1038/s41467-018-04049-3>.
40. Chamesian, A., Matsuda, M., Young, M., Wang, M., Zhang, Z.J., Liu, D., Tobin, B., Xu, Z.Z., Van de Ven, T., and Ji, R.R. (2019). Is optogenetic activation of Vglut1-positive A β low-threshold mechanoreceptors sufficient to induce tactile allodynia in mice after nerve injury? *J. Neurosci.* 39, 6202–6215. <https://doi.org/10.1523/JNEUROSCI.2064-18.2019>.
41. Warwick, C., Cassidy, C., Hachisuka, J., Wright, M.C., Baumbauer, K.M., Adelman, P.C., Lee, K.H., Smith, K.M., Sheahan, T.D., Ross, S.E., and Koerber, H.R. (2021). Mrgpr^{Cre} lineage neurons mediate optogenetic allodynia through an emergent polysynaptic circuit. *Pain* 162, 2120–2131. <https://doi.org/10.1097/j.pain.0000000000002227>.
42. Beaudry, H., Daou, I., Ase, A.R., Ribeiro-Da-Silva, A., and Séguéla, P. (2017). Distinct behavioral responses evoked by selective optogenetic stimulation of the major TRPV1+ and MrgD+ subsets of C-fibers. *Pain* 158, 2329–2339. <https://doi.org/10.1097/J.PAIN.0000000000001016>.
43. Schorscher-Petcu, A., Takács, F., and Browne, L.E. (2021). Scanned optogenetic control of mammalian somatosensory input to map input-specific behavioral outputs. *Elife* 10, e62026. <https://doi.org/10.7554/ELIFE.62026>.
44. Daou, I., Tuttle, A.H., Longo, G., Wieskopf, J.S., Bonin, R.P., Ase, A.R., Wood, J.N., De Koninck, Y., Ribeiro-da-Silva, A., Mogil, J.S., and Séguéla, P. (2013). Remote optogenetic activation and sensitization of pain pathways in freely moving mice. *J. Neurosci.* 33, 18631–18640. <https://doi.org/10.1523/JNEUROSCI.2424-13.2013>.
45. Iyer, S.M., Montgomery, K.L., Towne, C., Lee, S.Y., Ramakrishnan, C., Deisseroth, K., and Delp, S.L. (2014). Virally mediated optogenetic excitation and inhibition of pain in freely moving nontransgenic mice. *Nat. Biotechnol.* 32, 274–278. <https://doi.org/10.1038/NBT.2834>.
46. Abdo, H., Calvo-Enrique, L., Lopez, J.M., Song, J., Zhang, M.D., Usoskin, D., El Manira, A., Adameyko, I., Hjerling-Leffler, J., and Ernfor, P. (2019). Specialized cutaneous Schwann cells initiate pain sensation. *Science* 365, 695–699. <https://doi.org/10.1126/SCIENCE.AAX6452>.
47. Barik, A., Thompson, J.H., Seltzer, M., Ghitani, N., and Chesler, A.T. (2018). A brainstem-spinal circuit controlling nocifensive behavior. *Neuron* 100, 1491–1503.e3. <https://doi.org/10.1016/J.NEURON.2018.10.037>.
48. Sharif, B., Ase, A.R., Ribeiro-da-Silva, A., and Séguéla, P. (2020). Differential coding of itch and pain by a subpopulation of primary afferent neurons. *Neuron* 106, 940–951.e4. <https://doi.org/10.1016/J.NEURON.2020.03.021>.
49. Tashima, R., Koga, K., Sekine, M., Kanehisa, K., Kohro, Y., Tominaga, K., Matsushita, K., Tozaki-Saitoh, H., Fukazawa, Y., Inoue, K., et al. (2018). Optogenetic activation of non-nociceptive A β fibers induces neuropathic pain-like sensory and emotional behaviors after nerve injury in rats. *eNeuro* 5, ENEURO.0450-517.2018.2018. <https://doi.org/10.1523/ENEURO.0450-17.2018>.
50. Baumbauer, K.M., Deberry, J.J., Adelman, P.C., Miller, R.H., Hachisuka, J., Lee, K.H., Ross, S.E., Koerber, H.R., Davis, B.M., and Albers, K.M. (2015). Keratinocytes can modulate and directly initiate nociceptive responses. *Elife* 4, e09674. <https://doi.org/10.7554/ELIFE.09674>.
51. Zhou, X., Wang, L., Hasegawa, H., Amin, P., Han, B.X., Kaneko, S., He, Y., and Wang, F. (2010). Deletion of PIK3C3/Vps34 in sensory neurons causes rapid neurodegeneration by disrupting the endosomal but not the autophagic pathway. *Proc. Natl. Acad. Sci. USA* 107, 9424–9429. <https://doi.org/10.1073/PNAS.0914725107>.
52. Nassar, M.A., Levato, A., Stirling, L.C., and Wood, J.N. (2005). Neuropathic pain develops normally in mice lacking both Nav 1.7 and Nav 1.8. *Mol. Pain* 1, 24. <https://doi.org/10.1186/1744-8069-1-24>.
53. Agarwal, N., Offermanns, S., and Kuner, R. (2004). Conditional gene deletion in primary nociceptive neurons of trigeminal ganglia and dorsal root ganglia. *Genesis* 38, 122–129. <https://doi.org/10.1002/GENE.20010>.
54. Hires, S.A., Gutnisky, D.A., Yu, J., O'Connor, D.H., and Svoboda, K. (2015). Low-noise encoding of active touch by layer 4 in the somatosensory cortex. *Elife* 4, e06619. <https://doi.org/10.7554/eLife.06619>.
55. Kane, G.A., Lopes, G., Saunders, J.L., Mathis, A., and Mathis, M.W. (2020). Real-time, low-latency closed-loop feedback using markerless posture tracking. *Elife* 9, e61909. <https://doi.org/10.7554/ELIFE.61909>.
56. Kaupilla, T., Kontinen, V.K., and Pertovaara, A. (1998). Weight bearing of the limb as a confounding factor in assessment of mechanical allodynia in the rat. *Pain* 74, 55–59. [https://doi.org/10.1016/S0304-3959\(97\)00143-7](https://doi.org/10.1016/S0304-3959(97)00143-7).
57. Sorge, R.E., Martin, L.J., Isbester, K.A., Sotocinal, S.G., Rosen, S., Tuttle, A.H., Wieskopf, J.S., Acland, E.L., Dokova, A., Kadoura, B., et al. (2014). Olfactory exposure to males, including men, causes stress and related analgesia in rodents. *Nat. Methods* 11, 629–632. <https://doi.org/10.1038/nmeth.2935>.
58. Usoskin, D., Furlan, A., Islam, S., Abdo, H., Lönnnerberg, P., Lou, D., Hjerling-Leffler, J., Haeggström, J., Kharchenko, O., Kharchenko, P.V., et al. (2015). Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. *Nat. Neurosci.* 18, 145–153. <https://doi.org/10.1038/nn.3881>.
59. Prescott, S.A., Ma, Q., and De Koninck, Y. (2014). Normal and abnormal coding of somatosensory stimuli causing pain. *Nat. Neurosci.* 17, 183–191. <https://doi.org/10.1038/nn.3629>.
60. Mickle, A.D., Won, S.M., Noh, K.N., Yoon, J., Meacham, K.W., Xue, Y., McIlvried, L.A., Copits, B.A., Samineni, V.K., Crawford, K.E., et al. (2019). A wireless closed-loop system for optogenetic peripheral neuromodulation. *Nature* 565, 361–365. <https://doi.org/10.1038/S41586-018-0823-6>.
61. Iyer, S.M., Vesuna, S., Ramakrishnan, C., Huynh, K., Young, S., Berndt, A., Lee, S.Y., Gorini, C.J., Deisseroth, K., and Delp, S.L. (2016). Optogenetic

- and chemogenetic strategies for sustained inhibition of pain. *Sci. Rep.* 6, 30570. <https://doi.org/10.1038/SREP30570>.
62. Plaghki, L., Decruynaere, C., Van Dooren, P., and Le Bars, D. (2010). The fine tuning of pain thresholds: a sophisticated double alarm System. *PLoS One* 5, e10269. <https://doi.org/10.1371/JOURNAL.PONE.0010269>.
 63. Hsu, A.I., and Yttri, E.A. (2021). B-SOiD, an open-source unsupervised algorithm for identification and fast prediction of behaviors. *Nat. Commun.* 12, 5188. <https://doi.org/10.1038/S41467-021-25420-X>.
 64. Luxem, K., Mocellin, P., Fuhrmann, F., Kürsch, J., Miller, S.R., Palop, J.J., Remy, S., and Bauer, P. (2022). Identifying behavioral structure from deep variational embeddings of animal motion. *Commun. Biol.* 5, 1267. <https://doi.org/10.1038/S42003-022-04080-7>.
 65. Pereira, T.D., Tabris, N., Matsliah, A., Turner, D.M., Li, J., Ravindranath, S., Papadoyannis, E.S., Normand, E., Deutsch, D.S., Wang, Z.Y., et al. (2022). SLEAP: A deep learning system for multi-animal pose tracking. *Nat. Methods* 19, 486–495. <https://doi.org/10.1038/S41592-022-01426-1>.
 66. Graving, J.M., Chae, D., Naik, H., Li, L., Koger, B., Costelloe, B.R., and Couzin, I.D. (2019). DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. *Elife* 8, e47994. <https://doi.org/10.7554/ELIFE.47994>.
 67. Weinreb, C., Osman, M.A.M., Zhang, L., Lin, S., Pearl, J., Annapragada, S., Conlin, E., Gillis, W.F., Jay, M., Shaokai, Y., et al. (2023). Keypoint-MoSeq: parsing behavior by linking point tracking to pose dynamics. Preprint at bioRxiv. <https://doi.org/10.1101/2023.03.16.532307>.
 68. Bohoslav, J.P., Wimalasena, N.K., Clausing, K.J., Dai, Y.Y., Yarmolinsky, D.A., Cruz, T., Kashlan, A.D., Chiappe, M.E., Orefice, L.L., Woolf, C.J., and Harvey, C.D. (2021). DeepEthogram, a machine learning pipeline for supervised behavior classification from raw pixels. *Elife* 10, e63377. <https://doi.org/10.7554/ELIFE.63377>.
 69. Bohic, M., Pattison, L.A., Jhumka, Z.A., Rossi, H., Thackray, J.K., Ricci, M., Mossazghi, N., Foster, W., Ogundare, S., Twomey, C.R., et al. (2023). Mapping the neuroethological signatures of pain, analgesia, and recovery in mice. *Neuron* 111, 2811–2830.e8. <https://doi.org/10.1016/j.neuron.2023.06.008>.
 70. Zhang, Z., Roberson, D.P., Kotoda, M., Boivin, B., Bohoslav, J.P., González-Cano, R., Yarmolinsky, D.A., Turnes, B.L., Wimalasena, N.K., Neufeld, S.Q., et al. (2022). Automated preclinical detection of mechanical pain hypersensitivity and analgesia. *Pain* 163, 2326–2336. <https://doi.org/10.1097/J.PAIN.0000000000002680>.
 71. Blivis, D., Haspel, G., Mannes, P.Z., O'Donovan, M.J., and Iadarola, M.J. (2017). Identification of a novel spinal nociceptive-motor gate control for A δ pain stimuli in rats. *Elife* 6, e23584. <https://doi.org/10.7554/ELIFE.23584>.
 72. Ratté, S., Hong, S., DeSchutter, E., and Prescott, S.A. (2013). Impact of neuronal properties on network coding: Roles of spike initiation dynamics and robust synchrony transfer. *Neuron* 78, 758–772. <https://doi.org/10.1016/j.neuron.2013.05.030>.
 73. Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., et al. (2020). Array programming with NumPy. *Nature* 585, 357–362. <https://doi.org/10.1038/s41586-020-2649-2>.
 74. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* 17, 261–272. <https://doi.org/10.1038/s41592-019-0686-2>.
 75. McKinney, W. (2010). Data structures for statistical computing in Python. In *Proc. 9th Python in Science Conf.*, pp. 56–61.
 76. Hunter, J.D. (2007). Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* 9, 90–95. <https://doi.org/10.1109/MCSE.2007.55>.
 77. Bradski, G. (2000). The OpenCV Library25 (Dr. Dobb's Journal of Software Tools), pp. 120–125.
 78. Seabold, S., and Perktold, J. (2010). statsmodels: econometric and statistical modeling with Python. In *Proc. 9th Python in Science Conf.*, pp. 92–96.
 79. Waskom, M. (2021). seaborn: statistical data visualization. *J. Open Source Softw.* 6, 3021. <https://doi.org/10.21105/joss.03021>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms/strains		
Mouse: Ai32, Cre-dependent Chr2	JAX	RRID:IMSR_JAX:024109
MouseTRPV1 Cre	JAX	RRID:IMSR_JAX:017769
Mouse: Advillin Cre	Zhou et al. ⁵¹	N/A
Mouse: Nav1.8 Cre	Agarwal et al. ⁵³	N/A
Deposited data		
Data for model training	This paper	Mendeley Data: https://doi.org/10.17632/gn2wbkh7j3.1
Software and algorithms		
Computer code	This paper	GitHub: https://github.com/stofe95/ramalgo ; Zenodo: https://doi.org/10.5281/zenodo.10022925
Spike2	Cambridge Electronic Design	RRID:SCR_000903
Python 3.8	The Python Software Foundation	https://www.python.org/ ; RRID:SCR_008394
Matplotlib	The Matplotlib community	https://matplotlib.org/ ; RRID:SCR_008624
Seaborn	The Seaborn community	https://seaborn.pydata.org/ ; RRID:SCR_018132
Pandas	The Pandas community	https://pandas.pydata.org/ ; RRID:SCR_018214
Numpy	The Numpy community	https://numpy.org/ ; RRID:SCR_008633
Scipy	The Scipy community	https://scipy.org/ ; RRID:SCR_008058
OpenCV	The OpenCV community	https://opencv.org/ ; RRID:SCR_015526
VAME	Luxem et al. ⁶⁴	https://github.com/LINCellularNeuroscience/VAME ; RRID:SCR_022477
DeepLabCut	DeepLabCut developers Mathis et al. ³⁷	RRID:SCR_021391
DeepLabCut-live	DeepLabCut developers Kane et al. ⁵⁵	https://github.com/DeepLabCut/DeepLabCut-live
PyTrinamic	Tynamic Motion Control	https://github.com/trinamic/PyTrinamic

RESOURCE AVAILABILITY

Lead contact

Further inquiries or requests can be directed to Steve Prescott (steve.prescott@sickkids.ca).

Materials availability

This study did not generate new unique reagents. A full parts list is provided in [Table S1](#) which, together with details in [Figure S1](#), allows users to construct their own stimulator. The device will also be made commercially available in the near future.

Data and code availability

- All data reported in the paper are available upon request. Data used to train the mouse keypoint detection model using DeepLabCut is available at Mendeley Data (<https://doi.org/10.17632/gn2wbkh7j3.1>). This includes the full output directory from DeepLabCut, including 500 labeled frames used to train a Resnet-50 based model. Model performance depends on various factors (camera angle, lighting, background, etc.) and so performance will differ if those factors differ. Users should add their own training data representative of their testing conditions.
- Original code for running the device will be made available on GitHub upon publication (<https://github.com/stofe95/ramalgo>). An archival DOI is reported in the [key resources table](#).
- Any additional information required to reproduce the results reported in this study is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All procedures were approved by the Animal Care Committee at The Hospital for Sick Children (protocol #53451) and were conducted in accordance with guidelines from the Canadian Council on Animal Care. To express Chr2 selectively in different types of primary

somatosensory afferents, we used Ai32(RCL-ChR2(H134R)/EYFP) mice (JAX:024109), which express the H134R variant of ChR2 in cells expressing Cre recombinase. These were crossed with advillin^{Cre} mice (kindly provided by Fan Wang) to express ChR2 in all sensory afferents, TRPV1^{Cre} mice (JAX:017769) to express ChR2 in TRPV1-lineage neurons, or Nav1.8^{Cre} mice (kindly provided by Rohini Kuner) to express ChR2 in nociceptors. 8–16 week old male (n = 28) or female (n = 11) mice were acclimated to their testing chambers for 1 h on the day before the first day of testing, and each day for 1 h prior to the start of testing. Sex differences were not observed and data were therefore pooled.

METHOD DETAILS

Photostimulator

The stimulator is summarized in [Figure S1](#); A complete list of components (as numbered in [Figure S1](#), with part # and supplier information) is included as [Table S1](#). Briefly, collimated light from a red (625 nm) LED and blue (455 nm) LED is combined using a 550 nm cut-on dichroic mirror. Blue light is attenuated with a neutral density filter. This beam is combined with IR light from a 980 nm solid laser using a 900 nm cut-on dichroic mirror. The IR beam is expanded to fill the back of the focusing lens. The common light path is reflected upward with a mirror and focused to a spot 5 mm in diameter on the platform above. The surface area of the spot is $\sim 20 \text{ mm}^2$; photostimulus power values should be divided by this number to convert to light density (irradiance). Red light reflected off the mouse paw is collected by a photodetector through a 630 nm notch filter. All light sources are controlled by computer via appropriate drivers and a 1401 DAQ (Cambridge Electronic Design) using Spike2 (Cambridge Electronic Design) or custom software written in Python using Numpy,⁷³ Scipy,⁷⁴ Pandas,⁷⁵ Matplotlib,⁷⁶ and OpenCV.⁷⁷ The photodetector samples at 1 kHz with the same DAQ, thus synchronizing stimulation and withdrawal measurement. A camera provides video of the mouse from below (substage). Video is used for aiming with the help of visual feedback using the red light, which is turned on prior to photostimulation with blue or IR light. A near-IR light source is useful to improve lighting during high-speed video. In the manual version of the device, the device is slid by hand; leveling screws at the four corners of the breadboard have a plastic cap for smooth sliding. In the motorized version, the breadboard is attached to linear actuators (TBI Motion) via 3-D printed connectors. Motors are controlled via custom software. The user aims by keyboard or joystick, or fully automated aiming is left to a neural network trained to recognize the mouse paws.

Mechanostimulator

Computer-controlled mechanical stimulation was implemented using a 300C-I dual-mode indenter (Aurora Scientific). This stimulator can control and measure both force and length (height). Our software controls height in the same way LED/laser intensity is controlled for photostimulation. The exerted force is simultaneously measured at 1 kHz and recorded to computer. Because withdrawal is evident from changes in measured force, additional signals (e.g., reflectance, video) are not required for latency measurements.

Platform and enclosures

The platform and animal enclosures were custom made. Except when testing mechanical stimuli, the platform is 3 mm-thick clear Plexiglass mounted on 20 × 20 mm aluminum rails, adjusted to the desired height above the stimulator. For mechanical stimulation, plexiglass was replaced with a metal grate, specifically a stainless steel cooling rack. Various enclosure designs were tested. In the final design, clear Plexiglass tubes (outer diameter = 65 mm, thickness = 2 mm) cut in 12.5 cm lengths were used in conjunction with opaque white 3-D printed cubicle. The same tube used to transfer a mouse from its home cage is placed on the platform vertically and slid into a cubicle for testing (see [Figure 1B](#)). A notch cut into the base of each tube allows the experimenter to deliver a food reward, to poke the mouse (to wake or orient it), or to clean feces or urine from the platform if required. Each cubicle is 3-D printed and contains internal magnets that allow cubicles to be easily combined. Keeping the mice at fixed distances from each other is important for automated testing, where the stimulator is automatically translated a fixed distance when testing consecutive mice. To view the mouse in profile during high-speed video, we used a narrow rectangular chamber with clear walls on the front and left side (with a notch under the latter) and opaque walls at the rear and right side. In some cases, a mirror was placed at a 45° angle near the left wall to simultaneously capture a front view of the mouse.

Comparison with handheld fiber optic

Volunteer testers were instructed to use a fiber optic (multimode fiber optic patch cable, 1000 μm diameter core, NA = 0.48, SMA endings attached to 455 nm fiber-couple LED, Thorlabs) to apply a photostimulus to an s170C photodiode attached to a PM100D optical power meter (Thorlabs). The same photostimulus power was used for all trials, by all testers. The photodiode was covered with a paw-shaped cutout and placed face down on the plexiglass platform to simulate aiming at a real paw standing on the platform. The PM100D output was connected to a Power1401 data acquisition interface (Cambridge Electronic Design), sampling at 1 kHz. The Power1401 was also used to deliver command voltages to the LEDD1B LED driver (Thorlabs).

Automated withdrawal detection and latency measurement

Paw withdrawal is detected and its latency measured from the red reflectance signal using custom code written in Python. Red light is initiated prior to photostimulation with blue or IR light. Baseline reflectance is measured over the 0.5 s epoch preceding photostimulus onset. A running average across a 27 ms-wide window was used to remove noise. Withdrawal latency was defined as time

elapsed from photostimulus onset until the reflectance signal dropped below a threshold defined as 2 mV below baseline; the signal needed to remain below threshold for >20 ms to qualify as a response, but latency was calculated based on the start of that period. The 2 mV threshold value was chosen based on pilot experiments and then applied unchanged in all subsequent testing. Latencies thus extracted from the reflectance signal were compared to latency values extracted from high-speed video of the same withdrawal. In the latter case, paw height was extracted from video (see below) using DeepLabCut; latency was taken as the time taken for paw height to rise 6 pixels above baseline, defined as the mean height over the 0.5 s epoch preceding photostimulus onset. All latency measurements reported in the manuscript are based on automated reflectance-based measurements unless otherwise indicated.

High-speed video

High-speed video was collected with a Chronos 1.4 camera (Krontech) using a Computar 12.5–75 mm f/1.2 lens sampling at 1000 fps. To synchronize video with stimulation, the camera was triggered with digital pulses sent from the DAQ. Videos were compressed using H.264. Video was analyzed using DeepLabCut³⁷ to label the hind paw in sample frames and train a deep neural network to recognize the paw. This returned paw trajectories which were analyzed using custom code written in Python.

Pose estimation

DeepLabCut-Live was used to track mouse pose from substage video. While we stimulated only the left hindpaw, networks were trained to recognize the snout, front paws, hind paws, and tail base. The extra keypoints were intended to force the network to assume weights that would represent orientation well, and distinguish between the left and right paws. To train the neural network, we collected one video with 9 mice on the photostimulator platform and panned the camera around under the mice using the linear actuators, collecting 9 min of video. 500 frames were labeled and 95% were used for training a ResNet-50-based neural network with default parameters for 200,000 iterations. We validated on one shuffle and found a test error of 17.41 pixels (px) and train error of 2.62 px. These error values represent multiple keypoints; test error specifically related to the target hind paw is much lower (3.33 px). The image size was 640x480. Importantly, hind paws were not labeled when the paws were turned in a guarding position. This meant that the paws would not be recognized unless placed flat on the platform, and that stimulation would only occur when the paws were correctly oriented. Training was done on a 32 GB NVIDIA Tesla V100 GPU, while live inference for aiming was done on a 3 GB NVIDIA Quadro K4000 or NVIDIA GeForce RTX 4070 Ti (see below).

Different networks were required for different applications. To analyze paw withdrawal height, a separate neural network was trained using high-speed video of the mice in profile. DeepLabCut was used with the same parameters as above, training on 580 frames of high-speed video with a 1008x500 resolution. Test error = 5.05 px; train error = 2.32 px. For automated mechanical stimulation, another neural network was trained that could recognize the mouse on a metal grate. Again, the same parameters were used for DeepLabCut, but training on 100 frames with a 1280x800 resolution. Test error = 19.15 px; train error = 1.82 px. To validate photostimulus reliability with automated aiming, a network was trained to recognize a paw-shaped cut out covering a photodiode (see above). The same parameters were used as mentioned above, training on 200 frames with a 640x480 resolution, while labeling the center of the paw-shaped cutout. Test error = 2.43 px; train error = 2.1 px.

Automated aiming

The substage camera was aligned with the linear actuators such that movements in the x- and y-directions on video could be independently controlled by x- and y-linear actuators, respectively. The camera was also positioned such that the center of the frame was aligned with the photostimulation zone. x- and y-error signals were then calculated by taking distances from the DeepLabCut-live-based pose estimates for the target paw to the center of the frame (see Figure 7B for example frames). Inference for one frame on an NVIDIA GeForce RTX 4070 Ti could be completed in <20 ms, which is less than the duration of each frame for standard-rate video (1/30 fps = 33 ms). The x- and y-linear actuators were then independently driven with signals that were proportional to the error signals. Importantly, proportionality of the actuator speed to the error signal reduced translation speed as the target was approached, preventing overshoot. Once the target was centered in the frame within a tolerance of 3 px, a timer begins for a user-defined period (2 s for the data shown in Data S1) before stimulation is initiated. This delay assures that the mouse is immobile when initiating stimulation. If the target paw moves during the pre-stimulus interval, aiming is re-initiated and, once aligned, the timer is restarted. A separate timer can be set to stop this sequence and move to the next mouse when a certain mouse is too active to stimulate reliably.

Behavior extraction

Substage video was saved and compressed using H.264. DeepLabCut was used to identify the nose, left fore paw, right fore paw, left hind paw, right hind paw, and tail base. These key points were then fed into the VAME framework⁶⁴ using default parameters and 10 clusters to extract complex behaviors.

Capsaicin and heat hypersensitivity

A 0.5% w/v solution of capsaicin in mineral oil was prepared. Mice were lightly anesthetized using isoflurane and 5 μ L of capsaicin was injected into the left hind paw. Mice were allowed to recover for 15 min from anesthesia before thermal sensitivity was re-assessed.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical testing was performed using Python 3.8 using SciPy 1.9.3,⁷⁴ Statsmodels 0.13.2,⁷⁸ or with SigmaPlot v11. T-tests were used to compare means, Kolmogorov-Smirnov tests were used to compare distributions, and chi-square tests were used to test differences in frequency of observed behaviors. Visualizations were made with Matplotlib,⁷⁶ Seaborn,⁷⁹ and SigmaPlot.