Standardizing procedures to study sensitization of human spinal nociceptive processes: Comparing parameters for temporal summation of the nociceptive flexion reflex (TS-NFR)

Ellen L. Terry a, Christopher R. France b, Emily J. Bartley a, Jennifer L. DelVentura a, Kara L. Kerr a, Ashley L. Vincent a, Jamie L. Rhudy a,⁎

a The University of Tulsa, Department of Psychology, Tulsa, OK, United States
b Ohio University, Department of Psychology, Athens, OH, United States

Abstract

Temporal summation of pain (TS-pain) is the progressive increase in pain ratings during a series of noxious stimulations. TS-pain has been used to make inferences about sensitization of spinal nociceptive processes; however, pain report can be biased thereby leading to problems with this inference. Temporal summation of the nociceptive flexion reflex (TS-NFR, a physiological measure of spinal nociception) can potentially overcome report bias, but there have been few attempts (generally with small Ns) to standardize TS-NFR procedures. In this study, 50 healthy participants received 25 series of noxious electric stimulations to evoke TS-NFR and TS-pain. Goals were to: 1) determine the stimulation frequency that best elicits TS-NFR and reduces electromyogram (EMG) contamination from muscle tension, 2) determine the minimum number of stimulations per series before NFR summation asymptotes, 3) compare NFR definition intervals (90–150 ms vs. 70–150 ms post-stimulation), and 4) compare TS-pain and TS-NFR when different stimulation frequencies are used. Results indicated TS-NFR should be elicited by a series of three stimuli delivered at 2.0 Hz and TS-NFR should be defined from a 70–150 ms post-stimulation scoring interval. Unfortunately, EMG contamination from muscle tension was greatest during 2.0 Hz series. Discrepancies were noted between TS-NFR and TS-pain which raise concerns about using pain ratings to infer changes in spinal nociceptive processes. And finally, some individuals did not have reliable NFRs when the stimulation intensity was set at NFR threshold during TS-NFR testing; therefore, a higher intensity is needed. Implications of findings are discussed.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Some chronic pain may stem from an amplification of nociceptive signals within the central nervous system (i.e., central sensitization), even at the level of the spinal cord (Arendt-Nielsen and Graven-Nielsen, 2003; Bendtsen, 2000; Desmeules et al., 2003; Price and Staud, 2005; Verne and Price, 2002; Woolf, 1983). Thus, experimental methods to study these amplification processes could provide important insights into the initiation and maintenance of chronic pain. In animals, a series of noxious stimulations at a constant-intensity leads to a temporary hyperexcitability of nociceptive spinal cord dorsal horn neurons. This process has been called wind-up and is dependent on nociceptive C-fiber activation and NMDA receptors (Davies and Lodge, 1987; Mendell, 1966; Mendell and Wall, 1965), although repetitive stimulation of nociceptive A-delta fibers may also evoke it (Price et al., 1971). Wind-up may be one amplification process that contributes to central sensitization and chronic pain (Li et al., 1999; Woolf, 1983); however, wind-up is difficult to assess in humans given that it requires being able to directly record from dorsal horn neurons.

1.1. Using temporal summation of pain (TS-pain) to study wind-up in humans

Temporal summation of pain (TS-pain) is the progressive increase in pain ratings in response to a series of constant-intensity, noxious stimulations (Price, 1972; Price et al., 1977a). Furthermore, TS-pain is reduced by an NMDA-antagonist, indicating that it is dependent on NMDA receptors (e.g., Graven-Nielsen et al., 2000). Given these two parallels between wind-up and TS-pain, TS-pain is believed to be the psychophysical correlate of wind-up (Staud et al., 2006). However, there may be problems with this inference. Indeed, pain report can be biased, modulated, or otherwise influenced by factors that do not

influence spinal cord nociceptive neurons. For example, pain catastrophizing modulates pain ratings, but does not modulate spinal nociception (France et al., 2002a; France et al., 2004; Rhudy et al., 2009b; Rhudy et al., 2007). Thus, pain report does not always correlate with spinal nociceptive processes, so using pain report to assess wind-up can be problematic.

1.2. Nociceptive flexion reflex (NFR) is a physiological measure of spinal nociception

One method to assess sensitization of spinal nociception in humans that does not rely on subjective pain ratings is the nociceptive flexion reflex (NFR). The NFR is a protective withdrawal reflex evoked by brief electrocutaneous stimulations applied over the sural nerve at the ankle. If the electric stimulus is strong enough to drive nociceptive A-delta fibers, a lower limb withdrawal response is elicited that can be quantified from electromyogram (EMG) of the biceps femoris (hamstring) muscle. Because it requires the activation of A-delta fibers, it is a nociceptive response. Research has shown that the stimulus intensity that elicits the NFR (i.e., NFR threshold) is correlated highly with pain threshold, and the size/magnitude of the NFR correlates with pain ratings (Sandrini et al., 2005; Skljarevski and Ramadan, 2002). The NFR is used as a measure of spinal nociception because the reflex arc does not rely on supraspinal centers to be elicited (i.e., primary nociceptor → dorsal horn neurons → motoneurons), yet the reflex can be modulated from descending supraspinal controls (Rhudy et al., 2005; Sandrini et al., 2005). In sum, the NFR is a physiological measure of spinal nociceptive processes that can be used to avoid response bias which can contaminate subjective reports of pain (e.g., Aslaksen et al., 2007). For this reason, NFR should be a better measure of central (spinal cord) sensitization than pain ratings.

1.3. NFR can be used to assess sensitization of spinal nociceptive processes

Several observations suggest that the NFR can be used to assess sensitization of dorsal horn neurons (i.e., wind-up). First, as shown in Fig. 1, a series of constant-intensity electric stimulations (which depolarize nociceptive A-delta and C-fibers) evokes temporal summation of NFR (TS-NFR) as evidenced by a progressive increase in the magnitude of the NFR (Arendt-Nielsen et al., 1994; Arendt-Nielsen et al., 2000). Therefore, NFR demonstrates temporal summation. Second, TS-NFR appears to be dependent on NMDA receptors because an NMDA antagonist blocks TS-NFR (Arendt-Nielsen et al., 1995; Guiirimand et al., 2000). Given that wind-up is dependent on NMDA receptors, this indicates wind-up (hyperexcitability of dorsal horn neurons) and TS-NFR are mediated by a common physiological mechanism. Third, animal studies have simultaneously recorded the activity of single motor units (part of the circuitry associated with the motor output of the NFR) and the activity of wide dynamic range neurons in the dorsal horn of the spinal cord (Youn et al., 2003; You et al., 2004). When a series of noxious electrical stimulations was delivered, the studies by You et al. found that the temporal summation of the single motor units was correlated with the temporal summation of the wide dynamic range neurons, indicating the NFR can be used to assess temporal summation of spinal cord neurons. And finally, augmented TS-NFR has been noted in individuals with chronic pain who have central sensitization (e.g., Banic et al., 2004); specifically, the increase in NFR following a series of noxious stimulations is greater for those with chronic pain. Taken together, these lines of evidence suggest that TS-NFR can be used to assess sensitization of spinal cord neurons.

1.4. Prior studies attempting to promote a standardized methodology

Unfortunately, at this time, there are relatively few studies of TS-NFR compared to the number of TS-pain studies (e.g., Arendt-Nielsen et al., 1994; Farrell and Gibson, 2007; France et al., 2002b; Guirimand et al., 2000; Serrao et al., 2004), and comparisons across TS-NFR studies are hindered by the lack of a standardized TS-NFR protocol. Moreover, many of the TS-NFR studies that have attempted to promote a standardized methodology have used small samples (Ns<16), sometimes comprised exclusively of men (e.g., Arendt-Nielsen et al., 1994; Arendt-Nielsen et al., 2000; Guirimand et al., 2000).

We are aware of only three studies that compared different stimulus frequencies to determine the frequency that best evokes TS-NFR (Arendt-Nielsen et al., 1994; Arendt-Nielsen et al., 2000;
Farrell and Gibson, 2007). In the earliest study, Arendt-Nielsen et al. (1994) compared four stimulation frequencies (0.1, 1.0, 2.0, 3.0 Hz) in a sample of 8 men and found only the two highest frequencies (i.e., 2.0 and 3.0 Hz) evoked TS-NFR. One potential problem with this study is that these authors used a very large post-stimulation interval (70–200 ms post-stimulation) to quantify the NFR from the EMG signal. Quantification of the NFR is usually restricted to a narrow post-stimulus interval (e.g., 90–150 ms post-stimulus) because EMG activity prior to 90 ms can be associated with non-nociceptive A-beta fiber responses (also known as the RII reflex) and EMG activity after 150 ms can be due to voluntary and/or startle reactions (Arendt-Nielsen et al., 1994; Dowman, 1991, 1992; Sandrini et al., 2005). Therefore, the 70–200 ms post-stimulation interval used by Arendt-Nielsen et al. (1994) to quantify NFR could have contained non-nociceptive information unrelated to NFR. Additionally, when the 2.0 and 3.0 Hz stimulation frequencies were used, the authors noted increased muscle tension (baseline contamination) in the EMG signal that was due to carry-over from previous responses in the stimulation series.

This increased muscle tension could be problematic for several reasons. First, in order to correct for individual differences in resting muscle tension, NFR magnitude is typically defined as a change from pre-stimulation baseline activity (e.g., EMG activity in the 90–150 ms post-stimulation interval minus EMG activity in the 60 ms pre-stimulation baseline) (France et al., 2009; Rhudy and France, 2007). If the baseline EMG is contaminated by carry-over from prior responses, this poses a problem because the change score could underestimate the true NFR magnitude. To address this problem, some studies simply do not correct for baseline EMG activity (e.g., Arendt-Nielsen et al., 1994). While this partially deals with the problem, it makes comparing NFR magnitudes across individuals difficult because there is no correction for individual differences in resting EMG activity. A second problem is that the carry-over muscle tension could contaminate the post-stimulation interval used to quantify the NFR; thus, leading to an overestimation of the NFR magnitude. However, a third and potentially more important problem is that increased muscle tension can enhance NFR via processes unrelated to central sensitization (i.e., excitation of motoneurons) (Chan and Dallaire, 1989; Chan and Tsang, 1985; Desmedt and Godaux, 1976). So, if a participant begins to tense their leg during the stimulation series or if there is not enough time for the EMG signal to return to baseline levels before the next stimulation occurs, then the measurement of TS-NFR could be confounded. Thus, it is important to find a stimulation frequency that elicits TS-NFR and minimizes EMG carry-over (i.e., baseline contamination).

A subsequent study by Arendt-Nielsen et al. (2000) compared six stimulation frequencies (0.5, 1.0, 3.0, 5.0, 10.0, 20.0 Hz) in 13 men. They concluded that the highest stimulation frequencies (i.e., 10.0 and 20.0 Hz) led to the greatest temporal summation of NFR magnitude. Although this study used an 80–130 ms post-stimulation interval to quantify the NFR which helped exclude non-nociceptive responses, there are still a few problems that make the study difficult to interpret. First, they did not deliver the same number of stimuli for the different stimulus frequencies. Specifically, only two stimuli were delivered in the 0.5 Hz series, three stimuli were in the 1.0 Hz series, seven stimuli were in the 3.0 Hz series, eleven stimuli were in the 5.0 Hz series, 21 stimuli were delivered in the 10 Hz series, and 41 stimuli were delivered in the 20 Hz series. Because it is unclear how varying the number of stimuli in the series influences the results, the amount of temporal summation of the NFR elicited by the different stimulus frequencies cannot be compared. Second, the authors noted considerable baseline contamination (EMG carry-over) with the highest stimulation frequencies such that the EMG never returned to baseline levels in between stimulations (see Fig. 3 from Arendt-Nielsen et al., 2000). And third, it is unclear whether TS-NFR evoked by the higher stimulation frequencies (e.g., 10 Hz, 20 Hz) reflects temporal summation of nociceptive dorsal horn neurons (wind-up), because wind-up is only observed in response to stimulation frequencies ranging from 0.3 to 2.0 Hz (Eide, 2000).

In the largest of the three studies, Farrell and Gibson (2007) compared five stimulation frequencies (0.2, 0.25, 0.33, 1.0, 2.0 Hz) in a sample of 30 men and women. They noted that only the 2.0 Hz frequency evoked TS-NFR. Unfortunately, this study did not provide enough details to determine whether baseline contamination (EMG carry-over) was a problem. Further, they quantified NFR from a 70 to 130 ms post-stimulation interval. As previously noted, EMG activity as late as 150 ms post-stimulation can be associated with the NFR. So, it is not clear whether Farrell and Gibson's inability to observe summation of the NFR from frequencies slower than 2.0 Hz was due to their method of quantifying NFR. For this reason, future studies should examine different intervals for quantifying TS-NFR.

1.5. The present study

To determine the stimulation frequency that evokes TS-NFR and minimizes baseline EMG contamination (i.e., EMG carry-over), the present study presented 25 series of five noxious electrocutaneous stimulations. The stimulations were delivered in 5 Blocks (Fig. 2). Each Block contained five series, four of them were delivered at frequencies known to elicit wind-up in animals (0.33, 0.5, 1.0, and 2.0 Hz). Additionally, each block contained a series with a variable frequency. The variable frequency was designed to eliminate baseline contamination and used an “intelligent” stimulus delivery approach in which the computer monitored the biceps femoris EMG in between stimulations in “real time” and delivered the next stimulus in the series only after EMG activity returned to resting baseline levels. Specifically, the computer tried to deliver the next stimulus in the series 0.5 s after the first stimulus (i.e., at 2.0 Hz), but if EMG exceeded resting baseline then the stimulation would be delayed for up to 3.0 s (i.e., 0.33 Hz). The order of the different frequencies was randomized within Blocks and across participants.

To elicit TS-NFR, repeated noxious stimulations have to be delivered which can be difficult for participants to endure. Thus, a second goal of the study was to determine the minimum number of stimulations per series that are needed to elicit the maximum increase in NFR (i.e., the stimulation in which the increase in NFR asymptotes). Arendt-Nielsen et al. (1994) reported that NFR magnitude asymptotes between the 3rd and 5th stimulus in the series. However, they did not use statistical tests to come to this conclusion and the observation was based on data from only 3 male participants. Guirmand et al. (2000) reported that the NFR asymptotes between the 5th and 8th stimulation, but they also used a small sample of only 6 men. When a slightly larger sample of 9 males and 6 females was studied, Andersen et al. (2005) found the NFR asymptotes between the 3rd and 5th stimulus in the series. Thus, in an attempt to minimize the

<table>
<thead>
<tr>
<th>Phase 1</th>
<th>Phase 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFR Threshold</td>
<td>Pain Threshold</td>
</tr>
<tr>
<td>(order counterbalanced)</td>
<td></td>
</tr>
<tr>
<td>Block 1</td>
<td>Block 2</td>
</tr>
</tbody>
</table>

**Block**
- 5 series of noxious electrocutaneous stimulations presented
- Each series contained 5 noxious stimulations delivered at different frequencies (0.33 Hz, 0.5 Hz, 1.0 Hz, 2.0 Hz, Variable)
- Order of stimulation frequencies randomized
- NFR magnitudes quantified from each stimulation
- Pain ratings to each stimulation made at the end of each series

Fig. 2. Experimental protocol.

number of stimuli delivered to participants, we also determined the number of stimulations needed for NFR magnitude to asymptote.

A third goal was to compare intervals used to quantify the NFR. To our knowledge, no study to date has determined whether varying the interval influences the degree of temporal summation of NFR. Studies have used a number of different intervals to define NFR, e.g., 70–130 ms post-stimulation (Farrell and Gibson, 2007), 70–200 ms post-stimulation (Arendt-Nielsen et al., 1994), 80–130 ms post-stimulation (Arendt-Nielsen et al., 2000), 90–180 ms post-stimulation (Guirmand et al., 2000), but none have systematically compared them in the same group of participants. We have previously encouraged the use of a 90–150 ms post-stimulation interval (France et al., 2009; Rhudy and France, 2007) to avoid contamination by non-nociceptive responses that can occur earlier than 90 ms and later than 150 ms (Downman, 1992; Willer, 1977). While this 90–150 ms post-stimulation interval works well for responses to single stimulations, Arendt-Nielsen and colleagues have observed that repeated noxious stimulations that evoke TS-NFR may also shorten the onset latency of the NFR (Arendt-Nielsen et al., 1994). To determine whether the interval used to quantify NFR influences the degree of TS-NFR, the present study compared the 90–150 ms post-stimulation interval to an interval with an earlier onset (i.e., 70–150 ms post-stimulation). This change should account for a shortened NFR latency without being contaminated by the non-nociceptive RII reflex which typically occurs 40–60 ms post-stimulation (Hugon, 1973; Sandrini et al., 2005).

A fourth ancillary goal of this study was to examine TS-pain in response to different stimulation frequencies. Of the three studies that compared different stimulus frequencies to elicit TS-NFR (Arendt-Nielsen et al., 1994; Arendt-Nielsen et al., 2000; Farrell and Gibson, 2007), none assessed pain ratings to individual stimuli in the series. Thus, it is impossible to compare their findings to the TS-pain literature which always asks participants to report on each stimulation in the series.

2. Materials and methods

2.1. Participants

Participants were healthy individuals recruited from the University of Tulsa psychology subject pool and members of the Tulsa, OK community. Subject pool participants received research credit and community participants received a $25.00 honorarium. Participants were excluded for the following self-reported conditions: neurologic, cardiovascular, or circulatory problems; chronic pain; recent medical, cardiovascular, or circulatory problems; chronic pain; recent psychological trauma; use of over-the-counter pain medication within 24 h, or prescription pain medication within 2 weeks of participation; use of antidepressant, anxiolytic, or high blood pressure medications; having a body mass index of 35 or above (due to potential difficulties obtaining an NFR in individuals with high adiposity); and being under the age of 18.

Seventy participants consented to participate, but 12 were found to not meet inclusion criteria and 8 did not complete testing (1 reached tolerance before NFR threshold testing, 7 withdrew during temporal summation testing). Thus, data from 50 participants were available for analysis. There were no group differences in sex, ethnicity, age, or years of education between the 7 individuals who withdrew and the 50 who completed testing; however, NFR threshold was significantly higher for the participants that withdrew (M = 24.11 mA, SD = 8.16) versus those who completed (M = 14.40 mA, SD = 8.27; p = 0.005).

The majority of completers were female (62%, n = 31), white non-Hispanic (94%, n = 47), and employed fewer than 40 h a week (54%, n = 27). Completers had an average of 14 years of education (SD = 1.4) and averaged 21 years of age (SD = 2.8). All participants provided verbal and written informed consent. All participants were informed that they could withdraw from the study at any time. The study was approved by the Institutional Review Board (IRB) at The University of Tulsa.

2.2. Apparatus

Stimulus presentation, self-report ratings, and physiological data collection were controlled by a computer using a PC with dual monitor capacity, A/D board (PCI-6071E; National Instruments, Austin, TX), and LabVIEW software (National Instruments). One computer monitor was used by the experimenter to monitor physiological signals, and a second monitor was used by the participant to complete electronic questionnaires and to make ratings of electric stimuli. Testing was completed in a sound-attenuated and electrically-shielded testing chamber. Participants were monitored from an adjacent control room via a video camera connected to a flat panel television. Participants wore sound-attenuating headphones that allowed them to hear the experimenter’s instructions and could speak to the experimenter via the microphone on the video camera.

Electric stimuli were generated by a Digitimer stimulator (DS7; Hertfordshire, England) and delivered using a bipolar surface stimulating electrode (Nicolet, Madison, WI; 30 mm inter-electrode distance) attached to the left leg over the retromalleolar pathway of the sural nerve. A computer controlled the timing of the stimulations, and the maximum stimulation intensity was set at 50 mA. Physiological signals were amplified and filtered online using Grass Technologies (West Warwick, RI) Model 15LT amplifiers (with AC Modules 15A54 and DC Modules 15A12). Each signal was sampled at 1000 Hz.

The NFR was assessed from biceps femoris electromyogram (EMG) recorded from two active Ag–AgCl electrodes placed 10 cm superior to the popliteal fossa. A ground electrode was placed over the lateral epicondyle of the femur. Before electrodes were applied, the skin was cleaned with alcohol and exfoliated using an abrasive paste (Nuprep; Weaver and Company, Aurora, CO) to reduce impedances below 5 kΩ. All electrodes were then attached with self-adhesive collars after conductive gel (EC60; Grass Technologies) was applied.

2.3. Assessment of nociceptive responses

To assess pain sensitivity and TS-NFR, electric stimuli were delivered over the retromalleolar pathway of the sural nerve. Each electric stimulus was a train of 5 rectangular wave pulses of 1-ms duration with an inter-pulse interval of 3–ms (250 Hz). Because these 5-pulse trains are extremely brief (i.e., only 17 ms long), they are perceived as a single stimulation by the participant. For NFR threshold and pain threshold testing, single trains of 5 pulses were presented. For temporal summation testing, these 5-pulse trains were delivered as a series of 5 stimulations which were perceived as a series of 5 discrete stimulations. The stimulus intensity used during temporal summation testing was set at the higher of NFR threshold or pain threshold; therefore, NFR threshold and pain threshold was assessed prior to temporal summation testing.

2.3.1. Pain ratings

Pain intensity ratings were assessed from a computer-presented, 101-point scale ranging from 0 (no pain) to 100 (the most intense pain imaginable). Similar to our previous research (e.g., France et al., 2009; Rhudy et al., 2008; Rhudy and France, 2007; Rhudy et al., 2009a), a rating of 50 was labeled “painful” on the scale and corresponded to pain threshold. Participants were provided clear instructions on how to use the pain scale and were told that ratings below 50 would correspond to non-painful sensations and that a rating of 50 would correspond to a stimulus that just barely evoked pain. Anchoring pain at 50 ensured that the scale range for non-painful sensations and painful sensations were equivalent in length. To make a rating, participants used a computer mouse to drag a digital indicator vertically along the computerized scale and then pressed a
“submit” button to record their rating onto the computer. During NFR and pain threshold testing, participants provided a single pain rating after each stimulus, whereas during temporal summation testing they provided five pain ratings after each series of stimuli (i.e., 5 scales were presented at once and participants provided a pain rating for each stimulus in the series).

2.3.2. NFR threshold assessment

NFR threshold was assessed using three ascending–descending staircases of electric stimuli. The first ascending–descending staircase started at 0 mA and increased in 2 mA steps until an NFR was detected. NFR was defined as a mean biceps femoris EMG response in the 90–150 ms post-stimulus interval that exceeded the mean biceps femoris EMG activity during the 60 ms pre-stimulus baseline interval by at least 1.4 standard deviations (France et al., 2009; Rhudy and France, 2007). After an NFR was obtained, the current was decreased in 1 mA steps until an NFR was no longer detected. The second and third ascending–descending staircases used 1 mA steps. The interval between electric stimulations varied randomly between 8 and 12 s to reduce predictability and habituation. The stimulus intensity (mA) of the 2 peaks and 2 troughs of the last two ascending–descending staircases were averaged and used to define NFR threshold.

2.3.3. Pain threshold assessment

Similar to assessment of NFR threshold, pain threshold was assessed using three ascending–descending staircases of electric stimuli with a varying inter-stimulus interval of 8–12 s. The first ascending–descending staircase started at 0 mA and increased in 4 mA steps until pain threshold was reached (rating ≥ 50 on the pain rating scale described in Section 2.3.12). The current was then decreased in 2 mA steps until the participant rated a stimulus as ≤ 40 on the pain rating scale. (For 12 participants the criteria used to stop a descending staircase was a rating of ≤ 25; however, to shorten the length of the assessment the criterion was changed to ≤ 40.) The second and third ascending–descending staircases continued with 2 mA steps. Pain threshold was defined as the average intensity (mA) of the 4 stimuli first rated above and below 50 on the last two ascending and descending staircases.

2.3.4. Temporal summation of pain and NFR

Temporal summation was assessed from 5 Blocks of stimuli. Each Block consisted of 5 series that differed in their frequency (inter-stimulus interval or ISI is also noted): 0.33 Hz (3.0 s ISI), 0.5 Hz (2.0 s ISI), 1.0 Hz (1.0 s ISI), 2.0 Hz (0.5 s ISI), or variable frequency (0.5 s to 3.0 s ISI). These frequencies were chosen from the range of frequencies known to evoke wind-up in animals (Eide, 2000), with the expectation that slower frequencies (i.e., 0.33 Hz, 0.5 Hz) or the variable frequency might help to reduce baseline contamination (EMG carry-over). Each stimulation frequency was only presented once per Block and the stimulation frequency was randomly ordered within each Block by the computer so that every participant received a different order. After each series of stimuli, a set of 5 computer-presented pain rating scales were administered at the same time on the computer screen (see Section 2.3.1). Participants were instructed to rate pain intensity for each of the 5 stimulations in the series and then click on a button to submit their answers before the next series was delivered. Participants were offered short (1–2 min) breaks between Blocks.

During stimulations with a variable frequency (range = 0.5 s to 3.0 s ISI), biceps femoris EMG was monitored in “real-time” by the computer to ensure a stimulus was not delivered while EMG was above a participant’s resting EMG (resting EMG was determined prior to testing). Specifically, the ISI was 0.5 s (2.0 Hz) unless biceps femoris EMG in the 60 ms prior to a stimulus exceeded the resting EMG level, in which case the computer delayed stimulus delivery until EMG was at or below the resting level (or the 3.0 s max ISI was met). An upper limit of 3.0 s (0.33 Hz) was used because this corresponds to the slowest frequency shown to elicit wind-up (Eide, 2000; Price et al., 1977). Thus, the variable Hz was an attempt to use an “intelligent” stimulus delivery method that waited for EMG to return to resting baseline values in order to minimize EMG carry-over and baseline contamination. The mean ISI for the variable frequency in the current study was 1.19 s (SD = 0.93).

2.4. Procedure

Participants were tested while sitting comfortably in a reclining chair with the foot rest extended (knee angle approximately 160°). Upon arrival to the laboratory, participants were provided a description of the experimental procedures before informed consent was obtained. Next, a health status questionnaire and a brief interview were administered to assess inclusion/exclusion criteria. Participants were then trained to use the computer-presented pain rating scales and instrumented with electrodes. A 30 minute acclimation period followed during which questionnaires unrelated to the present study goals were administered.

Participants were informed that the experiment contained two phases (Fig. 2). Phase 1 assessed subjective and physiological reactions to single electric stimulations (NFR threshold and pain threshold testing) and phase 2 assessed reactions to 5 electric stimulations in a series (temporal summation testing). The order of NFR threshold and pain threshold testing was counterbalanced across participants (sex distribution was kept equivalent across testing order). Phase 1 took 20–30 min. After a 5 minute break, temporal summation testing (phase 2) was conducted which lasted about 20 min. The stimulus intensity used during temporal summation testing was set at NFR threshold or pain threshold, whichever was higher. At study completion, participants were provided their compensation (research credit or honorarium).

2.5. Preliminary data screening/scoring and data analysis

For all biceps femoris EMG trials during temporal summation, a trained experimenter visually inspected the EMG waveform offline for errors. Specifically, each waveform was displayed by computer and the EMG signal was examined to determine whether there were recording problems in the EMG signal or EMG activity that was unrelated to electric stimuli (i.e., muscle activity prior to the first stimulus in the series). All trials from one participant had to be excluded because the amplifier saturated in response to artifact from the electric stimulus. Three additional trials from other participants were eliminated due to movement. For EMG waveforms without errors, the average EMG in the baseline interval (60 ms pre-stimulus) and the average EMG in the NFR intervals (90–150 ms post-stimulus and 70–150 ms post-stimulus) were calculated automatically by the computer. Then, data from similar stimulation frequencies were averaged across Blocks.

Visual inspection was also used to determine whether NFRs were reliably elicited during temporal summation testing. Similar to procedures used in our previous research on NFR standardization (France et al., 2009; Rhudy and France, 2007), each rectified EMG waveform was examined using a computer program that clearly delineated the five baseline intervals (60 ms pre-stimulation) and the five NFR intervals (90–150 ms post-stimulation) for each stimulus series. An NFR was said to be present if at least one sizable peak was present in the NFR interval that was greater than EMG activity in the first baseline interval. By contrast, a reflex was said to not be present if EMG activity in the NFR interval mimicked EMG activity in the first baseline interval. We defined participants as NFR non-responders if there were ≤ 25% reflexes across all trials. This visual inspection determined that NFRs were not reliably elicited in 19 participants,
with an average of only 18% of NFRs elicited for these non-responders. Therefore, NFR data from these individuals could not be analyzed. To probe differences between responder and non-responder groups, we found NFR threshold was significantly lower for non-responders ($M = 8.13 \text{ mA}, SD = 4.65$) than responders ($M = 18.24 \text{ mA}, SD = 7.66$), $t(48) = 5.81, p < 0.001$. Additionally, pain threshold was significantly lower for non-responders ($M = 6.63 \text{ mA}, SD = 4.42$) than responders ($M = 14.08 \text{ mA}, SD = 8.48$), $t(47) = 3.53, p = 0.001$. This meant that the stimulus intensity used during temporal summation testing was significantly lower for non-responders ($M = 9.52 \text{ mA}, SD = 4.77$) than responders ($M = 20.99 \text{ mA}, SD = 8.43$), $t(48) = 5.41, p < 0.001$; hence, stimulus intensity may have been too low to reliably evoke NFRs during temporal summation in these individuals. There were no group differences in background or demographic variables (sex, ethnicity, age, and years of education) between responders and non-responders, all $ps > 0.05$.

Unless otherwise noted, data were analyzed using a 5 (Stimulus Number: 1–5) × 5 (stimulation frequency: 0.33 Hz, 0.5 Hz, 1.0 Hz, 2.0 Hz, and variable) completely-within subects repeated measures ANOVA. Significance level was set at $p = 0.05$. In the event of a significant interaction, the main effects are not reported and the interaction is decomposed by testing the simple effect of Stimulus Number using an $F$-test that compared the 5 means in the simple effect. If the $F$ was significant, then 7 post-hoc comparisons were conducted (Stimulation 1 vs. Stimulation 2, Stimulation 1 vs. Stimulation 3, Stimulation 1 vs. Stimulation 4, Stimulation 1 vs. Stimulation 5, Stimulation 2 vs. Stimulation 3, Stimulation 3 vs. Stimulation 4, Stimulation 4 vs. Stimulation 5) with a Bonferroni-adjusted alpha level set at $0.007$ ($\alpha = 0.05/7$ comparisons = 0.007). If the assumption of sphericity was not met, the Greenhouse–Geisser correction was made and the epsilon ($\varepsilon$) was reported. Partial eta squared ($\eta^2$) was reported as a measure of effect size for $F$-tests.

### 3. Results

#### 3.1. Baseline biceps femoris EMG

A significant Stimulation Frequency × Stimulus Number interaction was found, $F(16, 464) = 7.85, p = 0.004, \varepsilon = 0.09, \eta^2 = 0.21$ (Fig. 3A). $F$-tests for the simple effect tests for Stimulus Number were significant for all Stimulation Frequencies ($ps < 0.05$). However, none of the mean comparisons were significant for 0.33 Hz after the Bonferroni adjustments were made (i.e., no $ps < 0.007$). Together, these results indicate that 0.33 Hz was least affected by increases in baseline EMG, although increases in baseline muscle tension were generally a problem for all stimulation frequencies. In sum, results indicated that baseline EMG was contaminated (carry-over effects) for all stimulation frequencies, but contamination was most prominent for the 2.0 Hz stimulations (Fig. 3).

#### 3.2. Temporal summation of nociceptive flexion reflex (TS-NFR)

Given the observed baseline EMG results, NFR magnitude was calculated in three ways based on different methods to control for individual differences in baseline EMG:

1. **Unadjusted NFR** = mean EMG in the NFR interval (i.e., no control for baseline EMG)
2. **Locally Adjusted NFR** = mean EMG in the NFR interval minus mean EMG in the baseline interval (i.e., control for carry-over EMG contamination)

---

**Fig. 3.** Panel A represents baseline biceps femoris EMG activity depicted by the stimulus number in the 5 stimulus series and stimulation frequency (Hz). The baseline biceps femoris EMG is defined as the 60 ms activity prior to each electric stimulation. Baseline biceps femoris EMG increased across stimuli in the series regardless of the stimulation frequency. Panel B illustrates one participant’s biceps femoris EMG following a series of electrocutaneous stimulations at 2.0 Hz (from Block 1) with baseline EMG contamination (x-axis depicts 2500 ms trial). As can be seen, there is temporal summation of NFR (gradual increase in post-stimulation biceps femoris EMG). But, there is also a gradual increase in EMG activity in the pre-stimulation intervals (noted by the arrows).

3) Stimulation 1 Adjusted NFR = mean EMG in the NFR interval minus mean EMG from baseline interval prior to Stimulation 1 (i.e., control for resting EMG).

Because we were interested in whether the NFR interval influenced TS-NFR, each of the above scores was calculated using a 90–150 ms post-stimulation interval and a 70–150 ms post-stimulation interval. In general, as illustrated in Fig. 4, results of these analyses described in Sections 3.2.1, 3.2.2, and 3.2.3 below indicated that: 1) 2.0 Hz stimulations produced the largest and most consistent temporal summation of NFR (Fig. 4) regardless of the NFR interval used or the method used to adjust for baseline EMG, 2) quantifying the NFR from a 70 to 150 ms post-stimulation interval produced larger temporal summation of NFR than the 90–150 ms post-stimulation interval, and 3) temporal summation of NFR tended to asymptote by the 3rd stimulation in the series.

3.2.1. TS-NFR using 90–150 ms post-stimulus interval

Analysis of the unadjusted NFR magnitude revealed a significant Stimulation Frequency × Stimulus Number interaction, $F(16, 464) = 10.37, p = 0.001, \epsilon = 0.13, \eta^2 = 0.26$ (Fig. 4A). F-tests for the simple effects of Stimulus Number (i.e., temporal summation) indicated that unadjusted NFR magnitude differed across stimulations delivered at 1.0 Hz ($p = 0.04$) and 2.0 Hz ($p = 0.01$), but not 0.33 Hz ($p = 0.49$), 0.5 Hz ($p = .41$), or variable frequency ($p = 0.09$). Bonferroni-adjusted mean comparisons for 1.0 Hz indicated that there was a significant increase in NFR magnitude from Stimulations 1 to 4 and Stimulations 1 to 5. For 2.0 Hz, there was a significant increase in NFR magnitude from Stimulations 1 to 3, Stimulations 1 to 4, Stimulations 1 to 5, and Stimulations 2 to 3.

Analysis of the locally-adjusted NFR magnitude revealed a significant Stimulation Frequency × Stimulus Number interaction, $F(16, 464) = 7.03, p = 0.001, \epsilon = 0.15, \eta^2 = 0.20$ (Fig. 4B). F-tests for...
the simple effects of Stimulus Number indicated that locally-adjusted NFR magnitude differed across stimulations delivered at 2.0 Hz \((p = 0.03)\) and variable frequency \((p = 0.03)\), but not 0.33 Hz \((p = 0.22)\), 0.5 Hz \((p = 0.50)\), or 1.0 Hz \((p = 0.23)\). Bonferroni-adjusted mean comparisons for 2.0 Hz indicated that there was a significant increase in NFR magnitude from Stimulations 1 to 4 and Stimulations 1 to 5. For variable frequency, there was only a significant increase in NFR magnitude from Stimulations 1 to 3.

Analysis of the Stimulation 1 adjusted NFR magnitude revealed a significant Stimulation Frequency \(\times\) Stimulus Number interaction, \(F(16, 464) = 10.37, p < 0.001, \eta^2 = 0.26\) (Fig. 4C). F-tests for the simple effects of Stimulus Number indicated that Stimulation 1 adjusted NFR magnitude differed across stimulations delivered at 1.0 Hz \((p = 0.04)\) and 2.0 Hz \((p = 0.01)\), but not 0.33 Hz \((p = 0.49)\), 0.5 Hz \((p = 0.41)\), or variable frequency \((p = 0.09)\). Bonferroni-adjusted mean comparisons for 1.0 Hz indicated that there was a significant increase in NFR magnitude from Stimulations 1 to 4 and Stimulations 1 to 5. For 2.0 Hz, there was a significant increase in NFR magnitude from Stimulations 1 to 3, Stimulations 1 to 4, Stimulations 1 to 5, and Stimulations 2 to 3.

### 3.2.2. TS-NFR using 70–150 ms post-stimulation interval

Analysis of the unadjusted NFR magnitude revealed a significant Stimulation Frequency \(\times\) Stimulus Number interaction, \(F(16, 464) = 12.37, p < 0.001, \epsilon = 0.11, \eta^2 = 0.30\) (Fig. 4D). F-tests for the simple effects of Stimulus Number indicated that unadjusted NFR magnitude differed across stimulations delivered at 2.0 Hz \((p = 0.003)\) and variable frequency \((p = 0.02)\), but not 0.33 Hz \((p = 0.55)\), 0.5 Hz \((p = 0.39)\), or 1.0 Hz \((p = 0.08)\). Bonferroni-adjusted mean comparisons for 2.0 Hz indicated that there was a significant increase in NFR magnitude from Stimulations 1 to 3, Stimulations 1 to 4, Stimulations 1 to 5, and Stimulations 2 to 3. For variable frequency, there was a significant increase in NFR magnitude from Stimulations 1 to 3, Stimulations 1 to 4, and Stimulations 1 to 5.

Analysis of the locally-adjusted NFR magnitude revealed a significant Stimulation Frequency \(\times\) Stimulus Number interaction, \(F(16, 464) = 9.53, p < 0.001, \epsilon = 0.12, \eta^2 = 0.25\) (Fig. 4E). F-tests for the simple effects of Stimulus Number indicated that locally-adjusted NFR magnitude differed across stimulations delivered at 2.0 Hz \((p = 0.006)\) and variable frequency \((p = 0.02)\), but not 0.33 Hz \((p = 0.33)\), 0.5 Hz \((p = 0.52)\), or 1.0 Hz \((p = 0.23)\). Bonferroni-adjusted mean comparisons for 2.0 Hz indicated that there was a significant increase in NFR magnitude from Stimulations 1 to 3, Stimulations 1 to 4, Stimulations 1 to 5, and Stimulations 2 to 3. For variable frequency, the only significant increase in NFR magnitude was from Stimulations 1 to 3.

Analysis of the Stimulation 1 adjusted NFR magnitude revealed a significant Stimulation Frequency \(\times\) Stimulus Number interaction, \(F(16, 464) = 12.37, p < 0.001, \epsilon = 0.11, \eta^2 = 0.30\) (Fig. 4F). F-tests for the simple effects of Stimulus Number indicated that locally-adjusted NFR magnitude differed across stimulations delivered at 2.0 Hz \((p = 0.003)\) and variable frequency \((p = 0.02)\), but not 0.33 Hz \((p = 0.55)\), 0.5 Hz \((p = 0.39)\), or 1.0 Hz \((p = 0.08)\). Bonferroni-adjusted mean comparisons for 2.0 Hz indicated that there was a significant increase in NFR magnitude from Stimulations 1 to 3, Stimulations 1 to 4, Stimulations 1 to 5, and Stimulations 2 to 3. For variable frequency, there was a significant increase in NFR magnitude from Stimulations 1 to 3, Stimulations 1 to 4, and Stimulations 1 to 5.

### 3.2.3. Comparison of 90–150 ms vs. 70–150 ms post-stimulation intervals

The analysis of TS-NFR data indicated the 2.0 Hz stimulation frequency produced the largest temporal summation of NFR (Fig. 4); therefore, we directly compared the two NFR intervals when 2.0 Hz stimulations were presented. For these analyses, 2 (NFR Interval: 90–150 ms vs. 70–150 ms) \(\times\) 5 (Stimulation Number: 1–5) ANOVAs were conducted for each method of adjusting for baseline EMG. Table 1 presents these data. A significant Stimulation Number interaction was found for unadjusted NFR magnitude \(F(4, 116) = 6.83, p = 0.008, \epsilon = 0.32, \eta^2 = 0.19\), locally-adjusted NFR magnitude \(F(4, 116) = 6.83, p = 0.008, \epsilon = 0.32, \eta^2 = 0.19\), and Stimulation 1 adjusted NFR magnitude \(F(4, 116) = 6.83, p = 0.008, \epsilon = 0.32, \eta^2 = 0.19\). In all analyses, NFR magnitude was larger for the 70–150 ms interval versus the 90–150 ms interval during Stimulations 3, 4, and 5, but smaller during Stimulation 1 (\(ps < 0.01\)).

<table>
<thead>
<tr>
<th>Stimulation number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted NFR ((\mu V))</td>
<td>2.55a</td>
<td>5.15a</td>
<td>8.29a</td>
<td>8.35a</td>
<td>9.38a</td>
</tr>
<tr>
<td>(0.40)</td>
<td>(1.31)</td>
<td>(1.96)</td>
<td>(1.85)</td>
<td>(1.91)</td>
<td></td>
</tr>
<tr>
<td>Locally-adjusted NFR ((\mu V))</td>
<td>2.73</td>
<td>4.88</td>
<td>7.14</td>
<td>7.38</td>
<td>8.16</td>
</tr>
<tr>
<td>(0.43)</td>
<td>(1.15)</td>
<td>(1.65)</td>
<td>(1.59)</td>
<td>(1.62)</td>
<td></td>
</tr>
</tbody>
</table>

Note. All data are from 2.0 Hz stimulations. Regardless of the NFR interval used, NFR magnitude was adjusted for individual differences in resting EMG by subtracting the pre-stimulation EMG from each stimulation (locally-adjusted NFR) or from the pre-stimulation EMG for Stimulation 1 only (Stimulation 1 adjusted NFR).

\(a\) Denotes significant comparison between 70–150 and 90–150 ms NFR intervals.

### 3.3. Exploratory analyses: the influence of baseline contamination on TS-NFR

Given that the 2.0 Hz stimulations produced the largest TS-NFR, but also produced the greatest baseline EMG contamination (carry-over), this suggests that the observed increase in NFR from 2.0 Hz stimuli could be due to voluntary muscle contraction or contamination of the NFR interval by EMG carry-over. In an attempt to address this potential issue, we visually inspected each 2.0 Hz trial and categorized participants based on whether they did or did not have baseline EMG contamination. Specifically, each rectified EMG trial was displayed by the computer with the 60 ms prestimulation baseline intervals clearly demarcated. If EMG activity in the baseline interval for Stimulation 1, then a trial was said to have EMG contamination. This resulted in 11 participants categorized as having no contamination in their trials (i.e., no EMG carry-over). Fig. 1 depicts a trial from one of these participants.

We then analyzed the baseline EMG data to verify that experimenters accurately categorized participants as having contamination or no contamination. A 2 (Group: Contamination vs. No Contamination) \(\times\) 5 (Stimulation Number) ANOVA was conducted. This analysis resulted in a significant Group \(\times\) Stimulation Number interaction, \(F(4, 112) = 6.14, p = 0.01, \epsilon = 0.33, \eta^2 = 0.18\). F-tests for the simple effect of Stimulation Number found that the Contamination Group had a significant increase in baseline EMG \((p = 0.002)\), but the No Contamination Group did not \((p > 0.99)\).

We then selected data only from the No Contamination Group and analyzed their 2.0 Hz data using locally-adjusted NFR magnitudes...
scored from 70 to 150 ms and 90–150 ms post-stimulus intervals. One-way within-subjects ANOVAs with Stimulus Number as the repeated measures variable were conducted. The effect of Stimulus Number was significant for the 70–150 ms interval \([F(4, 40) = 11.40, p = 0.001, \eta^2 = 0.44, \eta^2_p = 0.53]\) and the 90–150 ms interval \([F(4, 40) = 5.80, p = 0.01, \eta^2 = 0.44, \eta^2_p = 0.37]\). To guard against Type II errors given that the sample size was small and the effects sizes were large (see partial eta squared values), we chose not to use Bonferroni adjustments for these comparisons (Maxwell, 2004). Mean comparisons indicated that NFR magnitude increased from Stimulations 1 to 2, Stimulations 1 to 5, and Stimulations 2 to 3 for both NFR intervals. Additionally, there was a significant increase in NFR magnitude from Stimulations 1 to 2 and Stimulations 1 to 4 when the 70–150 ms interval was used. In sum, this subgroup analysis indicates that the increase in NFR magnitude observed during 2.0 Hz stimulations was not due to increased muscle tension and/or EMG contamination.

### 3.4. Temporal summation of pain ratings (TS-pain)

Table 2 provides the mean pain ratings reported for each of the five stimulations across the different stimulation frequencies. A significant Stimulation Frequency x Stimulus Number interaction was found, \(F(16, 480) = 2.78, p = 0.05, \epsilon = 0.17, \eta^2 = 0.09\). Simple effects of Stimulus Number indicated that TS-pain was observed for 0.33 Hz \((p = 0.010), 0.5 \text{ Hz} (p = 0.018), 1.0 \text{ Hz} (p = 0.005), \) and 2.0 Hz \((p = 0.005), \) but not variable Hz \((p = 0.06).\) Bonferroni-adjusted mean comparisons indicated there were no significant differences between pain ratings at 0.33 Hz, 0.5 Hz, or variable Hz \((p > 0.05).\) For 1.0 Hz there was a significant increase in pain ratings from Stimulations 1 to 4 and Stimulations 1 to 5, as well as an increase from Stimulations 2 to 3, and Stimulations 4 to 5 \((p < 0.007).\) For 2.0 Hz there was an increase from Stimulations 1 to 3, Stimulations 1 to 4, and Stimulations 1 to 5, as well as increases from Stimulations 2 to 3, Stimulations 3 to 4, and Stimulations 4 to 5 \(p < 0.007).\) Bonferroni-adjusted mean comparisons also indicated that Stimulation 1 at the 2.0 Hz frequency was rated higher than all other frequencies, except the 1.0 Hz frequencies, which causes facilitation of the NFR, an effect that is not due to increased muscle tension contamination from the NFR interval used to quantify NFR, thereby superficially inflating the estimate of the NFR magnitude. Third, and perhaps most important, muscle tension may cause facilitation of the NFR via processes unrelated to spinal cord sensitization. Specifically, increased muscle tension can influence the excitability of motoneurons which causes facilitation of the NFR, an effect that is not due to sensitization of dorsal horn neurons or temporal summation (Chan and Dallaire, 1989; Chan and Tsang, 1985; Desmedt and Godaux, 1976).

Given that baseline EMG was contaminated for all stimulus frequencies, we mathematically adjusted for baseline contamination in three ways and compared them. First, we calculated NFR without adjusting for baseline EMG as is sometimes done in the TS-NFR literature (e.g., Arendt-Nielsen et al., 1994; Arendt-Nielsen et al., 2000; Farrell and Gibson, 2007; Guirimand et al., 2000). This approach ignores any problems associated with EMG carryover and contamination.

Second, we calculated NFR by subtracting local baselines (baselines immediately preceding each NFR). This approach removes individual differences in resting EMG and potentially removes muscle tension contamination from the NFR interval (assuming there is an equal amount of muscle tension in the baseline interval and NFR interval). However, it ignores the potential problem associated with increased motoneuron excitability.

Third, we calculated NFR as a change from the first baseline (Andersen et al., 2005). This approach corrects for individual differences in resting EMG, but ignores EMG contamination in the NFR interval and the problem of motoneuron excitability. We used these three methods of adjusting for baseline EMG to calculate NFR and we did so for two different post-stimulation intervals used to quantify NFR (70–150 ms post-stimulation, 90–150 ms post-stimulation). This allowed us to compare different baseline adjustment methods and different methods of quantifying NFR.

**Table 2**

<table>
<thead>
<tr>
<th>Frequencies</th>
<th>Stimulation number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>0.33 Hz</td>
<td>55.96</td>
</tr>
<tr>
<td></td>
<td>(2.73)</td>
</tr>
<tr>
<td>0.5 Hz</td>
<td>57.87</td>
</tr>
<tr>
<td></td>
<td>(2.90)</td>
</tr>
<tr>
<td>1.0 Hz</td>
<td>58.54</td>
</tr>
<tr>
<td></td>
<td>(3.13)</td>
</tr>
<tr>
<td>2.0 Hz</td>
<td>63.47</td>
</tr>
<tr>
<td></td>
<td>(3.25)</td>
</tr>
<tr>
<td>Variable</td>
<td>60.77</td>
</tr>
<tr>
<td></td>
<td>(2.71)</td>
</tr>
</tbody>
</table>

a Comparison with Stimulation 1 significant at \(p < 0.007\) (Bonferroni adjustment).
b Comparison between Stimulations 2 and 3 significant at \(p < 0.007\) (Bonferroni adjustment).
c Comparison between Stimulations 3 and 4 significant at \(p < 0.007\) (Bonferroni adjustment).
d Comparison between Stimulations 2 and 3 significant at \(p < 0.007\) (Bonferroni adjustment).

This study was designed to promote a standardized protocol for eliciting temporal summation of NFR (TS-NFR) by comparing four different stimulation frequencies known to elicit wind-up in animals (0.33, 0.5, 1.0, 2.0 Hz) and an “intelligent” stimulus delivery method intended to minimize baseline EMG contamination (variable frequency). Additionally, we compared two different NFR quantification intervals (90–150 ms vs. 70–150 ms post-stimulation) and three methods of correcting for individual differences in resting EMG activity (no adjustment, local baseline adjustment, stimulus 1 baseline adjustment).

Of the stimulation frequencies compared, all of them except 0.33 Hz led to significant EMG contamination (i.e., EMG carry-over). Increased muscle tension was present in the EMG by Stimulation 3 and continued to increase (or was maintained) until the 5th stimulation (Fig. 3). Contamination was most pronounced for the 2.0 Hz stimulations and would have been statistically significant even at the slowest frequency of 0.33 Hz had Bonferroni corrections not been used. These findings suggest that it is difficult for participants to keep their legs relaxed in between stimuli during temporal summation testing.

EMG contamination poses a significant problem for the study of TS-NFR in at least three ways. First, contamination of baseline EMG will affect the way NFR magnitude is calculated. NFR is often defined as a change from baseline in order to control for individual differences in resting (baseline) EMG activity that can be caused by differences in mass, subcutaneous adipose tissue, etc. (Rhudy and France, 2007). If baseline EMG is contaminated by muscle tension, this will affect the NFR change score. Second, some muscle tension may carry over into the post-stimulation interval used to quantify NFR, thereby superficially inflating the estimate of the NFR magnitude. Third, and perhaps most important, muscle tension may cause facilitation of the NFR via processes unrelated to spinal cord sensitization. Specifically, increased muscle tension can influence the excitability of motoneurons which causes facilitation of the NFR, an effect that is not due to sensitization of dorsal horn neurons or temporal summation (Chan and Dallaire, 1989; Chan and Tsang, 1985; Desmedt and Godaux, 1976).

Given that baseline EMG was contaminated for all stimulus frequencies, we mathematically adjusted for baseline contamination in three ways and compared them. First, we calculated NFR without adjusting for baseline EMG as is sometimes done in the TS-NFR literature (e.g., Arendt-Nielsen et al., 1994; Arendt-Nielsen et al., 2000; Farrell and Gibson, 2007; Guirimand et al., 2000). This approach ignores any problems associated with EMG carryover and contamination.

Second, we calculated NFR by subtracting local baselines (baselines immediately preceding each NFR). This approach removes individual differences in resting EMG and potentially removes muscle tension contamination from the NFR interval (assuming there is an equal amount of muscle tension in the baseline interval and NFR interval). However, it ignores the potential problem associated with increased motoneuron excitability.

Third, we calculated NFR as a change from the first baseline (Andersen et al., 2005). This approach corrects for individual differences in resting EMG, but ignores EMG contamination in the NFR interval and the problem of motoneuron excitability. We used these three methods of adjusting for baseline EMG to calculate NFR and we did so for two different post-stimulation intervals used to quantify NFR (70–150 ms post-stimulation, 90–150 ms post-stimulation). This allowed us to compare different baseline adjustment methods and different methods of quantifying NFR.
Regardless of the method used to adjust for baseline EMG or the NFR interval used, results revealed that stimulations delivered at 2.0 Hz elicited temporal summation of NFR that reached an asymptote by the 3rd stimulus in the series. Although 1.0 Hz and variable frequency stimulations led to temporal summation of NFR on some trials, summation was not consistently observed. Thus, consistent with prior research (Arendt-Nielsen et al., 1994; Arendt-Nielsen et al., 1995; Farrell and Gibson, 2007), our findings suggest the 2.0 Hz stimulations, but not slower frequencies, evoke TS-NFR. However, our study extends this prior research to show that the 70–150 ms post-stimulus interval produces a greater temporal summation of NFR than the 90–150 ms interval. Ostensibly, this is because the latency of NFR onset shortens with repetitive noxious stimulation. This notion is consistent with visual examination of the EMG signals (see Figs. 1 and 3B). However, we were unable to quantify the onset latency to empirically verify this due to problems with EMG contamination (carry-over) that made scoring the onset difficult (Fig. 3B). Given this, it appears that using a 70–150 ms post-stimulation interval to quantify TS-NFR is best.

Our study also extends prior research to suggest that 2.0 Hz stimulations evoke TS-NFR regardless of the baseline adjustment method. Not surprisingly, the degree of TS-NFR was smallest when stimulations were delivered at 2.0 Hz to evoke TS-NFR. However, as the current study and others addressed in some way. The most conservative approach would be to address this post-stimulation interval. Our laboratory is working on an alternative method to adjust for baseline contamination based on mathematical deconvolution, thereby increasing drop-out. Moreover, care should be taken to not make ratings to individual stimulations delivered at NFR threshold or pain threshold, whichever was higher. To our surprise, 19 of the 50 participants did not have reliable reflexes. We found that these individuals had significantly lower NFR thresholds and pain thresholds than NFR responders. As a result, we speculated that the stimulation intensity was set too low to evoke reliable NFRs during a series of stimuli. To test this hypothesis, we invited the 19 individuals back to the lab and retested them using a higher stimulation intensity (the higher of 150% NFR threshold or 150% pain threshold). Of the 9 retested, 7 had reliable reflexes. This preliminary data indicates that stimulation intensity should be set above NFR threshold during temporal summation testing. However, a higher stimulation intensity may make it more difficult for participants to tolerate the procedure thereby increasing drop-out. Moreover, care should be taken to not set the stimulation level too high, because that could cause a ceiling effect in the NFR such that increases in NFR magnitude cannot be observed.

In our laboratory we are testing a new procedure to determine the stimulation intensity used during temporal summation testing. A series of three stimulations are delivered at 2.0 Hz and the stimulus intensity for this 3-stimulation series is increased until an NFR is observed in response to the third stimulation in the series (responses to the first and second stimulations in the series are ignored). The “3-stimulation NFR threshold” intensity is then used to evoke reflexes during temporal summation testing. Preliminary observations suggest that this procedure produces reliable reflexes during temporal summation testing (Rhudy, unpublished observations).

4.2. NFR non-responders

In our study we initially set the stimulation intensity during temporal summation testing at NFR threshold or pain threshold, whichever was higher. To our surprise, 19 of the 50 participants did not have reliable reflexes. We found that these individuals had significantly lower NFR thresholds and pain thresholds than NFR responders. As a result, we speculated that the stimulation intensity was set too low to evoke reliable NFRs during a series of stimuli. To test this hypothesis, we invited the 19 individuals back to the lab and retested them using a higher stimulation intensity (the higher of 150% NFR threshold or 150% pain threshold). Of the 9 retested, 7 had reliable reflexes. This preliminary data indicates that stimulation intensity should be set above NFR threshold during temporal summation testing. However, a higher stimulation intensity may make it more difficult for participants to tolerate the procedure thereby increasing drop-out. Moreover, care should be taken to not set the stimulation level too high, because that could cause a ceiling effect in the NFR such that increases in NFR magnitude cannot be observed.

In our laboratory we are testing a new procedure to determine the stimulation intensity used during temporal summation testing. A series of three stimulations are delivered at 2.0 Hz and the stimulus intensity for this 3-stimulation series is increased until an NFR is observed in response to the third stimulation in the series (responses to the first and second stimulations in the series are ignored). The “3-stimulation NFR threshold” intensity is then used to evoke reflexes during temporal summation testing. Preliminary observations suggest that this procedure produces reliable reflexes during temporal summation testing (Rhudy, unpublished observations).

4.3. Implications and future directions

The results of the present study have a number of important implications. First, we recommend using a stimulation frequency of 2.0 Hz to evoke TS-NFR. However, as the current study and others have shown (Arendt-Nielsen et al., 1994; Arendt-Nielsen et al., 2000), contamination of EMG by muscle tension can be a problem. Therefore, the second implication is that baseline contamination should be addressed in some way. The most conservative approach would be to subtract local baselines. However, it is not clear whether this method removes some of the valid noceptive information from the NFR interval. Our laboratory is working on an alternative method to adjust for baseline contamination based on mathematical deconvolution,
which has been used successfully to adjust skin conductance responses to repetitive stimuli (Alexander et al., 2005). But, until a better method of handling the EMG contamination can be devised, we recommend that researchers use caution in interpreting the results from contaminated trials. Furthermore, the unadjusted NFR is not recommended when there are individual differences in resting biceps femoris EMG.

We found that both 90–150 ms and 70–150 post-stimulation intervals were adequate for quantifying NFR. However, the magnitude of TS-NFR was greater when the 70–150 ms interval was used. Therefore, the third implication is that the longer interval should be used to accommodate any shortening of NFR onset latency. A fourth implication is that a series of three stimulations appears to be adequate for studying TS-NFR. Our study observed that NFR reached an asymptote by the 3rd stimulus and failed to show further summation. While some prior studies found NFR reached an asymptote later (Arendt-Nielsen et al., 1994; Guirimand et al., 2000), these studies had significantly smaller sample sizes and were limited to men. So, it is not clear how well those studies generalize. Thus, we recommend defining TS-NFR as the difference between the stimulus 1 and stimulus 3. Reducing the number of stimuli in a series from five to three should produce TS-NFR and reduce the participant’s stimulus exposure by 40% thus reducing participant discomfort and potentially reducing drop-out (although drop-out was low in this sample at 14%). If researchers are also interested in recording pain ratings, then TS-pain can also be defined as the change from stimulus 1 to stimulus 3. However, it appears as though response bias is present in pain ratings when they are made at the end of the stimulus series; specifically, the higher the stimulation frequency, the higher the reported pain. Indeed, a fifth implication of our results is that pain ratings in response to stimulations at 2.0 Hz should not be used to make inferences about hyperexcitability of spinal cord neurons because of the response bias we observed. A sixth and final implication is that care should be taken when choosing a stimulation intensity to use during temporal summation testing. A stimulation intensity set too low may not produce reliable reflexes, whereas a stimulation intensity set too high may promote participant drop-out and/or NFR ceiling effects.

4.4. Limitations

While our study had a number of strengths, a few limitations should be noted. First, participants in the present study were young, healthy individuals; therefore, these results may not generalize to more diverse populations, including those with chronic pain. Second, while we did include both men and women in the study, analyses evaluating sex differences were not conducted due to the relatively lower number of male participants. Therefore, future research is needed to address potential sex differences in TS-NFR. Third, we did not move the stimulating electrode given that it is important to stimulate the sural nerve at the same location throughout testing (as the post-stimulation interval used to quantify NFR is associated with the conduction velocity of nociceptive A-delta fibers and moving the electrode could alter the distance to the spinal cord). Thus, it is possible that habituation to the electric stimulus could have influenced our findings. Fourth, the stimulation intensity used in the initial study did not appear to be high enough to reliably elicit NFRs in all participants during temporal summation testing. And fifth, as we noted above, our ability to accurately assess TS-pain was impaired by our use of retrospective pain ratings following the fifth stimulus in the series. However, given that the focus of the study was on standardizing procedures to evoke TS-NFR, we do not see this as a major limitation.

Despite these limitations, this study represents important progress toward determining the optimum parameters necessary to elicit TS-NFR. Ideally, development of a standardized methodology will help to promote the use of these techniques for the study of individual differences in pain processing and the study of processes that may contribute to the initiation and maintenance of chronic pain.

Acknowledgments

This research was partially supported by a grant from The University of Tulsa Office of Research and Sponsored Programs (awarded to Ellen Terry). The authors have no conflicts of interest to report.

References


