Empathy choice in physicians and non-physicians

C. Daryl Cameron¹* and Michael Inzlicht²

¹The Pennsylvania State University, University Park, Pennsylvania, USA
²University of Toronto, Ontario, Canada

Empathy in medical care has been one of the focal points in the debate over the bright and dark sides of empathy. Whereas physician empathy is sometimes considered necessary for better physician–patient interactions, and is often desired by patients, it also has been described as a potential risk for exhaustion among physicians who must cope with their professional demands of confronting acute and chronic suffering. The present study compared physicians against demographically matched non-physicians on a novel behavioural assessment of empathy, in which they choose between empathizing or remaining detached from suffering targets over a series of trials. Results revealed no statistical differences between physicians and non-physicians in their empathy avoidance, though physicians were descriptively more likely to choose empathy. Additionally, both groups were likely to perceive empathy as cognitively challenging, and perceived cognitive costs of empathy associated with empathy avoidance. Across groups, there were also no statistically significant differences in self-reported trait empathy measures and empathy-related motivations and beliefs. Overall, these results suggest that physicians and non-physicians were more similar than different in terms of their empathic choices and in their assessments of the costs and benefits of empathy for others.

Do physicians choose to feel empathy in everyday life? And should they do so? These intertwined questions seem to have an obvious answer. As members of the caring professions, one might argue that physicians ought to use empathy to understand and resonate with the internal experiences and feelings of their patients. Yet, empathy can often carry potent psychological and physical costs, particularly among people who are responsible for caring for the acute suffering of others. Given the real possibility of so-called burnout, perhaps physicians ought to refrain from too much empathy.

Motivational perspectives on empathy (e.g., Hodges & Biswas-Diener, 2007; Keysers & Gazzola, 2014; Zaki, 2014) suggest that empathy fluctuates as decision-makers balance the relative costs and benefits of empathy against alternative courses of action. Although empathizing with someone else may improve your relationship with that person, it may also be psychologically effortful and exhausting to empathize. These trade-offs may become particularly acute within medical contexts: The severity of outcomes for the patient can be a matter of life and death, meaning that the corresponding psychological costs for the care providers may be especially high. For example, professional care

*Correspondence should be addressed to C. Daryl Cameron, Department of Psychology, Rock Ethics Institute, University Park, PA 16802, USA (email: cdc49@psu.edu).

DOI: 10.1111/bjso.12342
workers who tend to humanize, or ascribe mental states to, their patients are more likely to show signs of emotional exhaustion and burnout (Vaes & Muratore, 2013). Such trade-offs involved in medical empathy have been discussed at length within psychology and health care (for reviews, see Gleichgerrcht & Decety, 2012; Halpern, 2007; Haque & Waytz, 2012).

Given these trade-offs, do physicians exhibit less empathy than non-medical controls? Some studies reveal that self-reported trait empathy declines during medical school and residency training (Hojat et al., 2009; Neumann et al., 2011; Smith, Norman, & Decety, 2017), yet some cross-sectional studies find that physicians are higher in trait empathy (e.g., Handford, Lemon, Grimm, & Vollmer-Conna, 2013), whereas other studies do not (e.g., Bellini & Shea, 2005). Moreover, some studies find variability across types of empathy measurement: For example, Smith et al. (2017) found declines in medical empathy during training on one self-report trait empathy measure (the Jefferson Scale of Physician Empathy; Hojat et al., 2001), but increases on a different self-report trait measure (the Questionnaire of Cognitive and Affective Empathy; Reniers, Corcoran, Drake, Shryane, & Vollm, 2011) and on a behavioural measure of mentalizing (Reading the Mind in the Eyes Test; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). Amidst these conflicting findings, empathic processes appear sensitive to motivational factors, as physicians who report increased satisfaction from helping others (i.e., ‘compassion satisfaction’; Stamm, 2009) exhibit increased empathy on trait measures (Gleichgerrcht & Decety, 2013) and increased empathic distress in response to depictions of pain (Gleichgerrcht & Decety, 2014).

One influential study examined differences in empathy between internal medicine physicians and non-physicians using a combination of physiology (i.e., event-related potentials) and self-report measures (Decety, Yang, & Cheng, 2010). Compared to non-physicians, physicians rated visual depictions of pain as less painful and unpleasant, showed less differentiation between painful and non-painful stimuli (e.g., hands being stuck with needles or brushed with Q-tips) in the N110 component (90–120 ms, thought to indicate spontaneous empathic resonance with pain) and in the P3 component (360–400 ms, thought to indicate cognitive evaluation), and showed less coupling between their subjective ratings of pain and their electrophysiological responses (Decety et al., 2010). Although there were no differences between groups on self-reported trait empathy measures, these results were interpreted as physicians strategically, and quickly, down-regulating empathy to avoid psychological costs (Decety et al., 2010). Although such results have been suggested to indicate empathy regulation, another possibility is that they reveal habituation: Work-related pain stimuli such as needles sticking hands may simply lose their emotional potency through exposure and repetition.

On the other hand, recent work used the same empathy for pain stimuli in a response interference paradigm, which instructed physicians and non-physicians to quickly judge painfulness of depicted target experiences while avoiding the influence of painful distracter experiences (Spring, Cameron, McKee, & Todd, 2019). Contrary to the implications of prior work, physicians and matched controls did not differ in their unintentional resonance with others’ pain depicted in distracter images. Moreover, physicians were more likely to intentionally empathize with others’ pain depicted in target images, suggesting a possible increase in goal-directed empathizing linked to medical experience (Spring et al., 2019).

The mixed evidence about physician empathy may be due to a number of factors. Existing studies have varied in which medical specializations they focus on, and in what time during medical training comparisons are being made. Yet aside from between-study
differences in the physicians being tested, another source of mixed findings may be due to how empathy is being assessed. Most prior work focuses on trait self-report measures of empathy, which require retrospective assessment of general patterns of empathizing across contexts (e.g., ‘how much concern do you tend to show in everyday life?’). Although this is a useful starting point for understanding medical empathy, it might not provide adequate resolution to capture how physicians manage their own feelings of empathy in real time.

There may also be discrepancies between how people self-assess their own empathy and how it is expressed behaviourally with others; for example, some recent work finds that physicians’ trait empathy on self-report measures does not correlate with patient ratings of physician empathy (Bernardo et al., 2018), suggesting that perhaps a greater focus on behavioural indicators of empathy might provide a clearer lens. How do physicians choose to structure their environments to facilitate or dampen empathy when confronted with the suffering of others? How do physicians make moment-to-moment decisions to empathize? And how does this compare to controls? To date, little work has directly addressed this question using behavioural assessments of empathy.

In the current work, we use a novel measure of empathy regulation to assess how physicians manage empathy when faced with suffering of others: the Empathy Selection Task (Cameron et al., 2019). In this task, participants complete a series of trials in which they are asked to freely choose between two card decks: One deck prompts them to engage in experience sharing with a social target, whereas the other deck prompts them to remain objectively detached from this target. This measure captures the spontaneous use of an emotion regulation strategy known as situation selection (Gross & Thompson, 2007): opting into or avoiding situations in which people expect to feel empathy. Situation selection is an effective strategy for managing emotional experiences because it allows people to shape their environments from precluding unwanted emotions from ever being elicited in the first place (Gross & Thompson, 2007) and has been studied in the context of empathy and prosocial behaviour (e.g., Shaw, Batson, & Todd, 1994). In addition to serving as a measure of empathy regulation, this approach can reveal motivational correlates. In post-task assessments of cognitive costs, participants rate each of the decks in the Empathy Selection Task on effort, aversion, and efficacy, which can then be examined in conjunction with empathy choice patterns (see Cameron et al., 2019). To the degree that physicians (vs. non-physicians) have different perceptions of cognitive costs of empathy due to their professional demands, this may shape their decisions of whether to engage in empathy.

Prior work developing this task attempted to understand the prevalence of empathy avoidance as well as its cause; this work revealed that empathy avoidance is not reducible to the avoidance of vicarious distress, as people avoid empathy for both negative and positive experiences (Cameron et al., 2019). Additionally, experimentally increasing feelings of efficacy for empathy reduced the empathy avoidance effect, suggesting that cognitive costs cause empathy avoidance, and are not simply a post hoc rationalization of empathy avoidance (Cameron et al., 2019). To address possible concerns that decks differ in mundane task features (e.g., boredom), some of the previous studies have structurally matched task features to require exactly the same kind of response while only varying amount of empathizing. For example, some task variants have asked participants to enter in exactly the same information (e.g., emotion keywords); some have participants choose between engaging in empathy for greater or less amounts of time (Cameron et al., 2019). Across these diverse operationalizations of the Empathy Selection Task, participants avoided empathy and this preference was clearly linked to perceptions of cognitive cost.
In the current study, we recruited physicians and demographically matched controls to complete the Empathy Selection Task. We examined whether physicians regulated empathy differently than demographically matched controls, and whether empathy choice was associated with measures of cognitive costs and individual differences. Specifically, we examined empathy in response to the suffering of child refugees, a context commonly studied in empathy research (e.g., Cameron & Payne, 2011). We opted to use this type of empathy context to maximize comparability to prior studies using the Empathy Selection Task (Cameron et al., 2019).

One concern with our use of child refugee stimuli might be that this context is not explicitly work-related (e.g., giving physicians the opportunity to have empathy for patients they expect to treat). However, many studies of physician empathy use context-general trait empathy measures (such as the Interpersonal Reactivity Index, which collects retrospective reports about general empathic tendencies across situations; see Bernardo et al., 2018; Gleichgerrcht & Decety, 2013). Such measures may be less able to tap into empathy across particular contexts of suffering and may have low resolution to capture how people actively regulate their empathy in different ways (for discussion, see Cameron, 2018). By contrast, here we present particular, rather than decontextualized, suffering contexts to test whether physicians and non-physicians actively choose to feel empathy. Other studies have used empathy stimuli that are more relevant to medical practice, such as body parts being stuck with needles (e.g., Decety et al., 2010). The empathy stimuli in the current study all depicted clear physical and emotional suffering, which is broadly relevant to the task of medical care, even if less directly so than body parts being stuck with needles in a medical setting. Although less work-related, the current stimuli have greater diversity of physical and emotional pain – and so are less susceptible to concerns about over-generalization from a particular pain context and type, or about emotional habituation to a mundane case. Additionally, in showing the complex range of physical and emotional symptoms of a full individual, rather than a decontextualized body part (e.g., a hand), the present stimuli might be more akin to contexts where physicians must appraise a patient’s symptoms. In summary, we selected a context that would have relevance for both physicians and non-physicians, be comparable to prior work on the Empathy Selection Task, and allow us to observe how physicians and non-physicians choose empathy in response to diverse depictions of pain.

Method
Participants
We recruited 130 participants using a Qualtrics Panel Survey, with 65 practicing physicians (34 female, 31 male, $M_{\text{age}} = 47.14$ years, $SD_{\text{age}} = 14.29$ years) and 65 non-physicians that were selected to be demographically matched (i.e., such that the control group also comprised 34 female, 31 male, with $M_{\text{age}} = 47.14$ years, $SD_{\text{age}} = 14.29$ years), with the groups also matched on education (having a doctoral/professional degree such as PhD, MD).¹ The panel recruitment was conducted by Qualtrics, such that the control participants were recruited to match on age, gender, and education level to the physicians. Sample size was determined to be as large as possible given the available

¹ One participant in this final sample was a replacement for an earlier participant, who was screened out for having entered inappropriate responses (i.e., indicating that they had copied and pasted responses, and making negative remarks about the study).
resources to fund panel recruitment with these populations. Sensitivity analyses with G*Power 3.1 suggest that given the sample size ($N = 130$), we have 80% power to detect a medium-sized effect ($d = .50$) of empathy choice between groups in a two-tailed independent samples t-test. Given that the average effect size in studies of social psychology is medium (Richard, Bond, & Stokes-Zoota, 2003), our study was limited to finding effect sizes that are slightly larger than the average in psychology.

**Materials and procedure**

Participants completed the measures listed below, in the order presented.

**Demographics**

At the beginning of the study, participants were asked ‘Are you a physician/doctor, currently practicing medicine?’ Based upon this screening, physicians then answered three questions: ‘How long ago (in hours) was your last shift in which you interacted with patients?’ ‘What is your medical specialty?’ ‘How long (in years) have you been a physician?’ All participants reported gender, age, race/ethnicity, political orientation (from *Extremely liberal* to *Extremely conservative* on a 7-point scale), and education level ($1 = \text{none}$, $2 = \text{elementary school}$, $3 = \text{some high school but no diploma}$, $4 = \text{High school (diploma or GED)}$, $5 = \text{trade school}$, $6 = \text{some college but no degree}$, $7 = \text{Associate’s degree (AA, AS, etc.)}$, $8 = \text{Bachelor’s degree (BS, BA, etc.)}$, $9 = \text{Master’s degree (MA, MPH, etc.)}$, $10 = \text{Doctoral/professional degree (PhD, MD, etc.)}$)

**Attention filter**

Participants then completed the first attention check, being instructed ‘Please enter the word ‘survey’ in the space below’. The participants included in the final panel successfully completed this filter.

**Empathy selection task**

At the beginning of the study, participants read the following instructions, adapted from previous uses of the Empathy Selection Task (Cameron *et al*., 2019):

In this task, you will complete a series of trials. On each trial, you will see two decks of cards: The deck on the left will always be labelled ‘DESCRIBE’ and the deck on the right will always be labelled ‘FEEL’. You should choose between these decks. Once you choose a deck, you will then see an image of a person. Depending on which deck you have chosen, you will be given one of two possible sets of instructions. If you choose from the deck labelled ‘DESCRIBE’, you will be told to be objective and focus on the external features and appearances of the person in the image. When completing this kind of trial, try to be as objective as possible. To be objective, do not let yourself get caught up in imagining what this person feels. On these trials, use one sentence to describe the age and gender of the person. If you choose from the deck labelled ‘FEEL’, you will be told to have empathy and focus on the internal feelings and experiences of the person in the image. When completing this kind of trial, try to feel as much empathy as possible. To be empathic, let yourself get caught up in imagining what this person feels. On these trials, use one sentence to describe the feelings and experiences of this person. You are free to choose from either deck on any trial, and should feel free to move from one deck to the other whenever you choose. If one deck begins to seem preferable, feel free to
choose that deck more often. Overall, this task will take the same amount of time regardless of which deck you choose.

If participants chose the FEEL deck they read: ‘Look at the person in the picture, and try to feel what this person is feeling. Empathically focus on the internal experiences and feelings of this person. Please write one sentence describing the experiences and feelings of this person’.

If participants chose the DESCRIBE deck they read: ‘Look at the person in the picture, and try to notice details about this person. Objectively focus on the external features and appearance of this person. Please write one sentence describing the age and gender of this person’.

Figure 1 displays a schematic of the Empathy Selection Task. To ensure that participants spent equivalent amounts of time on each trial, they could not submit their response until 5 s had elapsed. Participants completed 25 trials of the task. On each trial, participants viewed two card decks to choose between: the empathy deck was always a blue card deck on the right (labeled “FEEL”) and the objective deck was always a red card deck on the left (labeled “DESCRIBE”). The specific order of child refugee images was randomized across participants. The split-half reliability of the Empathy Selection Task (i.e., the correlation between proportions of empathy choice on the odd and even items) was high, \( r = .68 \).

Post-task assessment
After completing the Empathy Selection Task, participants provided open-ended responses to the prompts: ‘What was it like performing the task?’ ‘Did you develop a preference for one of the decks?’ and ‘How did you choose between the decks?’ These qualitative responses were not analysed for the study and are not reported further.

Figure 1. Schematic of Empathy Selection Task (cf. Cameron et al., 2019; Copyright © 2019 by the American Psychological Association, adapted with permission). This particular refugee image was not used in the study, but represents the kind of image that participants saw after making their choice between decks.
Cognitive costs
Cognitive costs were assessed for each deck (empathy, objectivity). For each deck, participants answered the following questions from the NASA Task Load Index (Hart & Staveland, 1988): ‘How mentally demanding was this deck?’ ‘How hard did you have to work to accomplish your level of performance with this deck?’ ‘How insecure, discouraged, irritated, stressed, and annoyed were you by this deck?’ ‘How successful were you in accomplishing what you were asked to do in this deck?’ For each deck, the first two questions were combined as effort, the third was aversion, and the fourth was efficacy.

Individual differences

State fatigue. Immediately after the NASA Task Load Index, participants also completed a single-item fatigue measure (Van Hooff, Geurts, Kompier, & Taris, 2007): ‘How fatigued do you currently feel?’ (from 1 = Not at all to 10 = Extremely). Unlike the NASA Task Load Index questions, which assessed the cognitive costs of each deck in the Empathy Selection Task, the single-item measure captured general levels of fatigue among participants completing the study.

Trait empathy. Participants completed two 7-item sub-scales from the Interpersonal Reactivity Index (Davis, 1983; from 1 = Does not describe me very well to 5 = Describes me very well): Empathic Concern (α = .81; e.g., ‘I often have tender, concerned feelings for people less fortunate than me’) and Perspective Taking (α = .70; e.g., ‘I try to look at everybody’s side of a disagreement before I make a decision’). There was an attention filter embedded within this questionnaire: ‘This is an attention filter. Please select “Describes me very well” for this statement’. All participants included in the final panel passed this attention filter.

Empathy-relevant motivations. Participants completed the Professional Quality of Life Scale (Stamm, 2009), which assessed thoughts and feelings in helping situations (from 1 = Never to 5 = Very often). The Professional Quality of Life Scale has three 10-item sub-scales including Compassion Satisfaction (α = .91; e.g., ‘I get satisfaction from being able to help people’), Burnout (α = .75; e.g., ‘Because of my work as a helper, I feel exhausted’), and Secondary Traumatic Stress (α = .88; e.g., ‘I am preoccupied with thoughts about the people I help’). One participant did not provide complete data for all Secondary Traumatic Stress items and was excluded for all analyses using that sub-scale.

Beliefs about empathy in medicine. Finally, participants completed two questions about the role of empathy in medicine (from 1 = Not at all to 5 = Very much): ‘How important is it that doctors have empathy for their patients?’ ‘Does empathy help doctors do their jobs well?’ Participants also provided an open-ended response to the question ‘In your opinion, what is the role of empathy in medical practice?’

Our data and syntax are available at the Open Science Framework: https://osf.io/ugvqt/.
Results

Empathy choice
We excluded data for six participants who skipped responses on the Empathy Selection Task, and one participant who entered an inappropriate response, leaving a final sample of 64 physicians and 59 non-physician controls. Table 1 displays empathy choice by participant group, and within physicians, split by medical specialty. Figure 2 displays violin plots of empathy choice by participant group. For our analyses, we first examined whether each group showed a preference to either avoid or approach empathy, using a one-sample $t$-test to compare the proportion of choosing the empathy deck across all trials against chance (as in Cameron et al., 2019). Controls displayed empathy avoidance, choosing the empathy deck 40.34% of the time ($SD = 22.05\%$), with empathy choice among this group deviating significantly below chance, $t(58) = -3.37, p = .001, 95\% CI$ of mean difference of choice from $0.50 [-0.15, -0.04]$, Hedges’ $g = -.43$. Physicians did not exhibit significant empathy avoidance, choosing the empathy deck 45.06% of the time ($SD = 22.84\%$), with empathy choice not deviating significantly from chance, $t (63) = -1.73, p = .089, 95\% CI [-0.11, -0.01]$, Hedges’ $g = -.21$.

Second, we used an ANOVA to compare empathy choice across the two participant groups. Critically, empathy choice did not differ between physicians and non-physicians, $F(1, 121) = 1.36, p = .246, 95\% CI = [-0.03, -0.13], \eta_p^2 = .01$, Hedges’ $g = .21$. Conducting a Bayesian analysis to quantify support for the null hypothesis of no difference in empathy choice between physicians and non-physicians resulted in anecdotal support for the null. Calculating Bayes factor in support of the null, with a default ‘objective’ prior using a zero-centred Cauchy distribution scaled at 0.707, resulted in $BF_{01} = 1.61$. That is, while there is more evidence for the null of no difference between

<table>
<thead>
<tr>
<th>Specialty</th>
<th>N</th>
<th>Proportion of emp. choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anaesthesiology</td>
<td>3</td>
<td>0.51 (0.24)</td>
</tr>
<tr>
<td>Cardiology</td>
<td>2</td>
<td>0.34 (0.37)</td>
</tr>
<tr>
<td>Emergency medicine</td>
<td>5</td>
<td>0.33 (0.20)</td>
</tr>
<tr>
<td>Family medicine</td>
<td>9</td>
<td>0.51 (0.25)</td>
</tr>
<tr>
<td>Gastroenterology</td>
<td>3</td>
<td>0.33 (0.23)</td>
</tr>
<tr>
<td>General</td>
<td>1</td>
<td>0.28 (-)</td>
</tr>
<tr>
<td>Hospitalist</td>
<td>3</td>
<td>0.52 (0.04)</td>
</tr>
<tr>
<td>Internal medicine</td>
<td>6</td>
<td>0.46 (0.36)</td>
</tr>
<tr>
<td>Nephrology</td>
<td>1</td>
<td>0.36 (-)</td>
</tr>
<tr>
<td>Neurology</td>
<td>1</td>
<td>0.04 (-)</td>
</tr>
<tr>
<td>Obstetrics</td>
<td>1</td>
<td>0.56 (-)</td>
</tr>
<tr>
<td>Oncology</td>
<td>1</td>
<td>0.64 (-)</td>
</tr>
<tr>
<td>Ophthalmology</td>
<td>2</td>
<td>0.42 (0.03)</td>
</tr>
<tr>
<td>Paediatrics</td>
<td>13</td>
<td>0.52 (0.22)</td>
</tr>
<tr>
<td>Preventive medicine</td>
<td>1</td>
<td>0.64 (-)</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>7</td>
<td>0.44 (0.23)</td>
</tr>
<tr>
<td>Radiology/imaging</td>
<td>2</td>
<td>0.22 (0.25)</td>
</tr>
<tr>
<td>Surgery</td>
<td>2</td>
<td>0.48 (0.11)</td>
</tr>
<tr>
<td>Vascular medicine</td>
<td>1</td>
<td>0.52 (-)</td>
</tr>
<tr>
<td>Physicians (combined)</td>
<td>64</td>
<td>0.45 (0.23)</td>
</tr>
<tr>
<td>Non-physicians</td>
<td>59</td>
<td>0.40 (0.22)</td>
</tr>
</tbody>
</table>
physicians and non-physicians than the alternative, the evidence in favour of the null is very small and anecdotal. More evidence is required to reach a judgement with more certainty.

Finally, we used an equivalence test procedure (TOST; Lakens, 2017) to test for whether the observed effect was practically worthwhile. The TOST procedure based on Welch’s $t$-test indicated that the observed effect size ($d = .21$) was not significantly within the equivalent bounds of $d = -.2$ and $d = .2$ (or in raw scores: $-0.04$ and $0.04$), $t(120.74) = 0.06, p = 0.523$. This non-significant result of the equivalence test means that we cannot reject the hypothesis that the true effect for differences in empathy choice between physicians and non-physicians is at least as strong as $d = .2$.

**Cognitive costs**

Next, we used repeated measures ANOVAs to examine whether participants rated the empathy and objective decks differently in terms of cognitive costs assessed by the NASA Task Load Index (i.e., effort, aversion, efficacy). Across physicians and non-physicians, participants rated the empathy (vs. objective) deck as more effortful, $F(1, 122) = 58.45, p < .001$, 95% CI [0.60, 1.02], $\eta^2_p = .32$, Hedges’ $g = .80$, more aversive, $F(1, 122) = 16.57, p < .001$, 95% CI [0.22, 0.64], $\eta^2_p = .12$, Hedges’ $g = .36$, and less efficacious, $F(1, 122) = 8.96, p = .003$, 95% CI [-0.49, -0.10], $\eta^2_p = .07$, Hedges’ $g = -.29$. When adding participant group into these models, there were no moderations of these differences (all $ps > .190$; all $BF_{01} > 2.48$), suggesting no differences for these empathy costs across physicians and non-physicians. Table 2 presents the difference score between decks for each of the three cognitive costs, separately for physicians and non-physicians. Figure 3 displays the deck differences (i.e., empathy deck minus objective deck) for each of the three cognitive costs.

We then used correlational analyses to examine whether deck differences in felt cognitive costs would relate to empathy choice in the Empathy Selection Task, as in
previous work (Cameron et al., 2019). Participants were less likely to choose the empathy (vs. objective) deck when they felt less efficacious completing the empathy (vs. objective) deck, \( r(121) = -0.31 \), \( p < 0.001 \), but there were no relationships with effort, \( r(121) = 0.06 \), \( p = 0.501 \), or aversion, \( r(121) = 0.04 \), \( p = 0.653 \). Similar to prior work (Cameron et al., 2019), efficacy at empathy was the strongest correlate of empathy choice, such that both physicians and non-physicians were more likely to choose empathy when they felt successful at generating it.

**Individual differences**

For the next analyses, we examined whether each of the individual differences (i.e., trait empathy; empathy-relevant motivations on the Professional Quality of Life Scale; self-reported fatigue and time since last shift; and beliefs about empathy in medicine) correlated with empathy choice. Additionally, to complement the group comparisons in empathy choice on the Empathy Selection Task, we conducted ANOVAs to compare groups on these individual differences.

![Violin plots depicting differences in NASA Task Load Index cognitive costs (effort, aversion, efficacy) for the empathy deck (Feel) minus the objective (Describe) deck, split by participant group. The median and quartiles are presented within each participant group.](image)

**Figure 3.** Violin plots depicting differences in NASA Task Load Index cognitive costs (effort, aversion, efficacy) for the empathy deck (Feel) minus the objective (Describe) deck, split by participant group. The median and quartiles are presented within each participant group.
Trait empathy
Consistent with some prior work showing no differences across physicians and controls on self-reported trait empathy (e.g., Decety et al., 2010), participant groups did not differ on the Interpersonal Reactivity Index sub-scales of Empathic Concern, $F(1, 121) = 0.39$, $p = .532$, $\eta^2_p = .00$, $BF_{01} = 4.37$, or Perspective Taking, $F(1, 121) = 0.27$, $p = .607$, $\eta^2_p = .00$, $BF_{01} = 4.60$. Table 2 displays group differences in all individual difference measures. Bayesian analyses, with default ‘objective’ priors, suggest modest support for the null, meaning there is modest evidence that the physicians and non-physicians controls have equivalent levels of Empathic Concern and Perspective Taking, as assessed by this measure. Empathy choice correlated positively with Empathic Concern, $r(121) = .19, p = .034$, and Perspective Taking, $r(121) = .18, p = .045$. Thus, participants who reported being prone to greater compassion and perspective taking in everyday life were more likely to choose empathy.

Professional quality of life
Participant groups did not differ on the Professional Quality of Life sub-scales of Compassion Satisfaction, $F(1, 121) = 2.88$, $p = .093$, $\eta^2_p = .02$, $BF_{01} = 1.42$; Burnout, $F(1, 121) = 2.52$, $p = .115$, $\eta^2_p = .02$, $BF_{01} = 1.66$; or Secondary Traumatic Stress, $F(1, 120) = 0.27$, $p = .603$, $\eta^2_p = .00$, $BF_{01} = 4.58$. Empathy choice correlated positively with Compassion Satisfaction, $r(121) = .26, p = .004$, and negatively with Burnout, $r(121) = -.19, p = .037$, but not Secondary Traumatic Stress, $r(120) = -.07, p = .420$. Thus, participants who derived more satisfaction from helping others were more likely to choose empathy, and participants who experienced more exhaustion from helping were less likely to choose empathy.

Fatigue and time since last shift
For the single-item fatigue measure, on average participants reported moderate fatigue ($M = 5.08$ on a 10-point scale, $SD = 2.11$). Participant groups did not differ on the single-item fatigue measure, $F(1, 121) = 2.08$, $p = .152$, $\eta^2_p = .02$, $BF_{01} = 2.04$, and fatigue did not correlate with empathy choice, $r(121) = -.01, p = .892$. Within the physicians group, three physicians did not indicate hours since last shift, and two physicians did not indicate number of years being a doctor. Empathy choice was uncorrelated with hours since last shift, $r(59) = .05, p = .706$, and number of years being a doctor, $r(61) = -.02, p = .892$.

Beliefs about empathy in medicine
The two belief items were averaged together ($r = .64, p < .001$) as an index of positive beliefs about empathy in medicine. One participant did not complete both items and was excluded from analysis. Participant groups did not differ in their beliefs about empathy in medicine, $F(1, 121) = 1.35$, $p = .247$, $\eta^2_p = .01$, $BF_{01} = 2.83$. Participants with more positive beliefs about empathy in medicine were more likely to choose empathy, $r(120) = .25, p = .005$. Beliefs that empathy is useful for medical practice associated with increased empathy choice, and this relationship was not unique to physicians.

Exploratory analyses of empathy choice by specialty
In addition to the primary analyses, we also conducted two exploratory analyses to examine how empathy choice among physicians might differ depending on factors
related to their medical specialties. First, we grouped the specialties based upon whether they are typically seen as ‘person-centered’ (i.e., focusing on interactions with patients) or ‘technology-centered’ (i.e., focusing on technical instrumentation), based upon classifications in previous studies of medical specialty choice (Taber, Hartung, & Borges, 2011; Yufit, Pollock, & Wasserman, 1969; see also Borges & Savickas, 2002). Person-centred specialties (group $n=46$) were coded as family medicine, general medicine, internal medicine, obstetrics, pediatrics, and psychiatry; additionally, we coded cardiology, gastroenterology, nephrology, neurology, oncology, and vascular medicine as sub-branches of internal medicine and thus as falling within this category. Technology-centred specialties (group $n=12$) were coded as anaesthesiology, emergency medicine, radiology/imaging, and surgery. The remaining physician specialties (group $n=6$; hospitalist, ophthalmology, and preventive medicine) were not clear relative to previous criteria and excluded from analysis. Empathy choice did not differ significantly across technology-centred specialties ($M=38.00\%, SD=21.20\%$) and person-centred specialties ($M=46.17\%, SD=24.36\%$), $F(1, 56)=1.13, p=.293, \eta^2_p=.02$, though we note that we had low power for this test.

Second, we grouped the medical specializations in Table 1 based upon independent ratings by practicing doctors of whether these fields posed high, medium, or low risks of emotional exhaustion (Gleichgerrcht & Decety, 2014, Table 1). Prior work finds that physicians in high-exhaustion specializations report higher empathic distress in response to video depictions of pain (Gleichgerrcht & Decety, 2014). We coded the high-exhaustion group ($n=28$) as including emergency medicine, oncology, pediatrics, psychiatry, and surgery; the medium-exhaustion group ($n=21$) as anaesthesiology, cardiology, gastroenterology, hospitalists, internal medicine, nephrology, neurology, obstetrics, and vascular medicine; and the low-exhaustion group ($n=15$) as family medicine, general medicine, ophthalmology, preventive medicine, and radiology. Empathy choice did not differ across the low-exhaustion ($M=45.07\%, SD=23.35\%$), medium-exhaustion ($M=42.86\%, SD=25.33\%$), and high-exhaustion groups ($M=46.71\%, SD=21.28\%$), $F(2, 61)=0.17, p=.847, \eta^2_p=.01$.

**Discussion**

The current study examined whether physicians, compared to control matched on gender, age, and education, were more or less likely to choose to feel empathy for the suffering of others. Overall, physicians did not show a clear preference for empathy or objectivity, with their choice preferences not deviating statistically from chance. While non-physicians showed a clear preference to avoid empathy, replicating previous work (Cameron *et al.*, 2019), there was no statistical difference between physicians and matched control non-physicians in empathy choice. We should note, however, that one limitation of the current study is that the sample was under-powered to detect a small effect of the size observed between groups here ($d=.21$), and although there was no statistical difference, we cannot rule out a small effect or assume that they are equivalent (Lakens, 2017).

The current work qualifies past work on physician empathy, particularly findings that physicians seem to exhibit diminished empathy on behavioural measures compared to matched controls (e.g., Decety *et al.*, 2010). In terms of how they regulated themselves in relation to empathic situations, doctors did not show a greater preference to avoid empathy than controls. If anything, our results show the opposite, albeit non-significant,
pattern. It may be that the emotion regulation strategy measured here—situation selection—is less likely to capture empathy differences across these groups, given that physicians have self-selected into professional situations in which empathy is often demanded. Instead, perhaps it is the immediate, unintentional empathic response, once in an empathic situation, that differentiates physicians and their non-medical counterparts (e.g., Decety et al., 2010). However, other work finds no such difference in unintentional empathy between physicians and controls in empathic situations, and that physicians are actually better at intentionally empathizing with others’ pain (Spring et al., 2019). The current results are consistent with the findings by Spring et al. (2019): Statistically speaking, physicians were not less likely to show empathy, and although there was low power to detect the small observed effect, physicians were descriptively more likely to choose empathy. Returning to the issue of career self-selection, the current study cannot fully speak to this question, as the consistent message across a diversity of empathy measures is that physicians are quite similar to non-physicians.

Consistent with these findings, physicians were no more or less likely than control participants to see empathy (vs. objectivity) as more effortful, aversive, or inefficacious, suggesting that, at least within the context of the Empathy Selection Task, felt cognitive costs did not differ across groups. Physicians and non-physicians both saw empathy as hard work, and to the degree that they viewed empathy in this way, chose it less (replicating Cameron et al., 2019). Self-reported trait empathy associated with increased empathy choice, providing convergent validity for the Empathy Selection Task, and furthermore did not itself differ between physicians and non-physicians.

Advancing previous work, here we see an arguably motivational measure of empathy at work associates with empathy choice: Positive experiences from helping others (the Compassion Satisfaction scale) associated with increased empathy choice, and self-reported negative physiological experiences from helping others (the Burnout scale) associated with reduced empathy choice. Similarly, believing that empathy is more relevant for helping others in medical contexts associated positively with empathy choice. These associations provide further construct validation of the Empathy Selection Task: Whereas empathy-relevant reward motivations (i.e., Compassion Satisfaction) associate positively with empathy choice, empathy-relevant avoidance motivations (i.e., Burnout) associate negatively with empathy choice. These results build on prior work on motivated empathy regulation (e.g., in the context of mass suffering, Cameron & Payne, 2011; or with stigmatized drug addicts, Cameron, Harris, & Payne, 2016) and extend these relationships to medically relevant empathy motivations.

**Novel directions in medical empathy**

One important conclusion from the current study is that physicians had equivalent levels of self-reported empathy, motivations to empathize, and belief that empathy is important for medicine as non-physicians. Bayesian analyses suggest that physicians and controls appeared to be, perhaps surprisingly, equivalent in terms of their empathy regulation behaviour and self-reported empathic experience, in their perceptions of psychological and professional costs of empathy, and in their beliefs about whether empathy sustains good medical practice. In the current study, there was much variation in specialization within the physician group, and it may be that future studies would be able to test more specific functional questions about empathy costs within particular specializations.

Within samples, physicians and non-physicians appeared similar, rather than different, in regard to their empathic profiles. Our study is not the first to find a lack of difference in
self-report trait measures of empathy, but it does contrast with prior work showing differences in physiological responding thought to reflect empathy (e.g., Decety et al., 2010). One advantage of our study for developing stronger theoretical inferences about physician empathy is that it captures empathy regulation per se – whereas some prior work suggests that empathy deficits in physicians reflect habitualized emotion regulation to avoid being over-aroused (Decety et al., 2010), such an interpretation is necessarily speculative because the regulation processes are not directly observed. Additionally, this earlier work cannot rule out that physicians may have habituated to a very particular and routinized context – that is, dealing with administering or witnessing needles sticking patients rather than regulating empathy per se, and such results might not generalize to empathy for more multi-faceted physical and emotional experiences.

By contrast, the Empathy Selection Task captures the emotion regulation strategy of situation selection over repeated instances. This increased focus on the process of empathy regulation itself may be one reason for discrepant findings from previous work. Other work mentioned in the introduction has focused on changes in physician empathy over time during medical training, rather than cross-sectional comparisons of physicians and non-physicians. Such work has tended to use self-report trait measures of empathy, which can capture emotion but may also reflect self-identity and perceived norms of the profession. The Empathy Selection Task is well suited to capture longitudinal changes in how physicians might relate to their own feelings of empathy over time, and in concurrence with possible changes in the felt cognitive costs of empathizing. For example, physicians could be assessed on the Empathy Selection Task after each year of medical training, and the perceived effort and efficacy of empathy could be assessed as well, to test whether greater experience with medical training might shape the cognitive work of empathy in a manner that facilitates future changes in empathy regulation.

Future work can build upon this initial study by including different variants of the Empathy Selection Task. One limitation of the current study is that the empathy stimuli were not defined as being directly relevant for medical practice. Considerations of the context of empathy may be important; some prior work has found that nurses with more work experience showed reduced valence and arousal ratings for pain depicted in a hospital context, but not in a home context (Cheng, Chen, & Decety, 2017). It may be that physicians and matched controls would show more differences in empathy choice if the task was contextualized to be about medical treatment. For instance, on each empathy trial participants might be asked not only to experience share with targets, but also to use that empathy in some manner to reach a diagnosis. Or, physicians might be provided with different expectations about forthcoming medical contexts – such as understanding the nature of a patient’s pain, or performing surgery – which might amplify or inhibit empathy choice, respectively. In other words, giving empathic choices functional utility within a medical context might provide a more sensitive test for profession-based differences in empathy choice (see also Tamir, 2009). Doing so might also increase interest in the task among physicians, and interest has been shown to motivate performance on even effortful tasks (Milyavskaya, Galla, Inzlicht, & Duckworth, 2018).

Prior work developing the Empathy Selection Task has found that people choose empathy more when they feel it is more valuable (Cameron et al., 2019), and in the current study, participants did generally endorse the belief that empathy is useful for medical practice, and such a belief was linked with choosing empathy more strongly. The child refugee stimuli in the current study presented clear physical and emotional suffering, which is broadly relevant for the task of medical care even if the need for such care was not explicitly stipulated in the instructions. It might be that physicians and non-
physicians are more likely to show differences in choices about compassion, rather than empathy, given that compassion might be more likely to associate with the kinds of caregiving behaviours seen in many medical contexts. Furthermore, it is an intriguing and unanswered empirical question whether empathy regulation that might occur in explicitly work-related diagnosis contexts might generalize to other, non-work contexts (i.e., if empathy avoidance becomes a habitualized strategy that carries over across situations).

Another limitation of the current work is that the sample had low power, with a sample size that was targeted to be as large as possible given available resources. Descriptively, the results indicate that physicians were slightly more likely to choose empathy compared to non-physicians, and the Bayesian analyses and equivalence test suggest the possibility of a small effect that may be more detectable in a larger sample. If physicians do choose empathy more often than non-physicians, then this raises interesting new questions about whether people who choose empathic emotions might self-select into medical practice, explaining the small effect, or whether people who undergo medical training change in their empathic choices. The current results contradict prior work suggesting that physicians spontaneously feel less empathy than non-physicians (Decety et al., 2010), as at least on a descriptive level, they suggest the opposite pattern. By better understanding when, why, and how physicians choose to intentionally engage with empathy, future research might broaden conclusions about the challenges of medical empathy.

The current work simultaneously highlights the bright and dark sides of empathy. On the one hand, participants felt that empathy for others was cognitive work – rating it as effortful, aversive, and inefficacious (Cameron et al., 2019). This finding in itself suggests a possible dark side to empathy that at least in the context of suffering strangers, physicians may find it cognitively taxing. Moreover, felt inefficacy at empathizing associated with choices to avoid empathy for physicians and non-physicians alike, suggesting a possible explanation for ‘dark side’ empathy deficits (e.g., Bloom, 2017). On the other hand, motivational measures in the current study associated with increased empathy choice across participant groups, including beliefs that empathy is useful for medicine and that compassion is satisfying. Such results hold promise for a bright side of empathy: To the degree that people choose empathy in different ways depending on their goals, then activating different goals may hold potential for changing how people regulate feelings of empathy – and in particular, those people such as physicians who are charged with providing care to others. For example, if physicians were led to reflect on how empathy may play an important role in medicine prior to making empathic choices, perhaps this might increase their tendency to choose empathy; similarly, if physicians were cued to recall prior experiences in which empathy for patients led to positive, satisfying outcomes, one might expect a similar result.

**Conclusion**

In summary, do physicians choose empathy, and should they do so? We find that physicians do not show a clear preference to approach or avoid empathy. Nevertheless, they do perceive empathy to be cognitively taxing, entailing effort, aversiveness, and feelings of inefficacy, and perceptions of inefficacy associated with reduced empathy choice. Physicians who derived more satisfaction and less burnout from helping were more likely to choose empathy, and so too if they believed that empathy is good, and useful, for medical practice.
More generally, in the current work, physicians did not show statistically meaningful differences from demographically matched controls in trait empathy, empathy regulation behaviour, motivations to approach or avoid empathy, or beliefs about empathy’s use for medicine. Although it has often been suggested that physicians exhibit different levels of empathy due to the demands of medical care, the current results suggest that physicians are much like everyone else, sensitive to the relevant costs and benefits of empathizing.

Acknowledgements

This research was funded by National Science Foundation grant #1660707 to Cameron and grant 435-2014-0556 from the Social Sciences and Humanities Research Council of Canada to Inzlicht. The first author also acknowledges support from the University of Iowa for this research. The first author presented this work at talks at DePaul University and Duke University.

References


Received 22 January 2019; revised version received 14 August 2019