Teachers perceive distinct competency profiles in soft and hard social robots for supporting learning

Luca M. Leisten[1], Nathan Caruana[2], Emily S. Cross[1]

***Abstract*— The promise of social robot applications for children's education has attracted growing enthusiasm over the past decade, with the potential to augment and support diverse learning outcomes. However, the adoption of education robots and their expected benefits for children are yet to be realised, due to complexity, cost, and variability between robots. Soft robots offer a possible solution. However, a concern is that these robots may be seen as less competent, decreasing their adoption and utility in learning environments. In this preregistered, mixed-methods study, we investigated teachers' (*n* = 120) perception of 12 hard and soft social robots along different dimensions, learning tasks, roles, and contexts. Teachers perceived hard robots as more competent, human-like, and familiar than soft robots. Soft robots were perceived as more physically/visually warm. Hard robots were also more likely to be perceived as suitable for "technical tasks" and adopting a teacher/tutor role for supporting the learning of adults or groups. Soft robots were more likely to be evaluated as suitable for use with younger learners in individual learning contexts and playing the role of a co-learner/novice. This study provides a detailed account of how soft and hard robot features influence teachers' perceptions of robot suitability for education applications. The findings directly inform how to optimise the design and situation of social robots to maximize adoption, effectiveness, and accessibility across diverse learners and learning contexts. By highlighting the nuanced trade-offs between competence and warmth, this research challenges theoretical assumptions that complex hard robots are universally superior in educational settings.**

I. INTRODUCTION AND RELATED WORK

Research on social robots in children's education has received sustained and increasing interest over the last two decades, with a variety of different robots and roles for these robots studied [1]–[3]. Despite growing research on the potential benefits of robots for learning, the first 15 years of this field has been characterized by opportunistic research, which has deployed very different robots to very different learning contexts to address very different research questions [1]. While this research has continued to highlight the promise of robots, it has not yet systematically examined which specific robot features (i.e., dimensions such as human-likeness, size, socialness, intelligence, functionality, interaction modalities, and shape; [2]) are likely to make robots acceptable and useful as education aids. This knowledge is critical for informing a human-centered approach to the design and implementation of effective social robots for children's education [4].

One particularly underexplored, yet potentially important, aspect of robot variability is their 'softness' (see **Fig. 1** for hard and soft robot examples). Soft robots are an entire category of robots on their own and are typically built from soft materials, such as silicon, rubber, fur or fabric, and sometimes resemble biological structures without rigid joints [5]. As soft robots can be made from a wide range of, often recycled, materials with openly available building plans and assets (e.g., [6]), they are generally far less costly, more sustainable, and accessible than their hard robot counterparts. Additionally, these types of robots are usually less technically complex. On the one hand, this might limit their usability for certain educational tasks, but on the other hand, this may increase accessibility for teachers, who are often responsible for deploying all education resources in the classroom [7].

Fig. 1. Examples of hard and soft social robots.

While soft robots could potentially make social robots more accessible, a recent study found that a robot's soft exterior (in this case fur) was associated with perceptions of decreased competence compared to harder exteriors, such as plastic or metal [8]. On the other hand, the degrees of freedom a robot displayed (i.e., the number of possible positions or motions a robot could execute) increased ratings of warmth (measured through ratings of how tolerant, warm, good natured, and sincere the robot is perceived as; [8]) and competence (measured through ratings of how competent, confident, independent, competitive, and intelligent the robot is perceived as; [8]), when evaluating these robots based on static pictures.

Perceived competence is an important factor to support children's learning [4], and previous work has shown that perceptions of robot competence and warmth influence robot acceptance [9], and thus possibly its effectiveness in an education context. How an embodied technology, such as a social robot, engages a learner on a cognitive level, and thereby possibly influences learning outcomes, largely depends on the assigned role of the robot (e.g., peer vs. teacher; [1]), the context in which the robot should be deployed [10], and the robot's perceived skills and appearance (e.g., physical human-likeness, intelligence, friendliness, competence; [11]).

It is crucial that the functional capabilities of robots be considered when evaluating their appropriateness for certain

[1]Social Brain Sciences Lab, ETH Zurich, 8092 Zurich, Switzerland `lleisten@ethz.ch, ecross@ethz.ch`
[2]Flinders University, 5042 Adelaide, Australia

applications. However, the appearance of robots is also critical, as it can affect early evaluations of a robot's unseen functional capabilities, role, and social characteristics [11], [12]. Such first impressions are often key in shaping use adoption [13] and robot acceptance [14]. One important distinction in the appearance of many commercially available robots is their softness. As such, it might affect people's first impression of a robot and consequently shape their expectations regarding its capabilities and application. As people's expectations are crucial for a successful application of technology [13], comprehensively examining robot feature profiles, social roles, context, and target users when evaluating their appropriateness as an education aid is important.

A. The present study

In this study, we sought to address how soft and hard robot features influence teachers' perceptions of robot suitability for different education applications. We collected survey data to comprehensively canvas teachers' perceptions of various soft and hard social robots and their suitability for different educational contexts, roles, and tasks. Teachers are critical stakeholders to involve in the design and deployment of education robots and should be considered in research on HRI in education contexts since they ultimately decide whether and how this technology is adopted in the classroom [7].

Prior work on the perception of robots based on their physical features has largely relied on depicting robots in static pictures [8], [11], [12]. While these studies show that two dimensional depictions of a robot's visual features reliably influence people's perceptions, it is also worth noting that important information regarding a robot's functionality (e.g., how it moves, interacts, or conveys emotion) is lost. In [4], almost half of participants changed their preference for a certain robot after seeing the robot in motion. Specifically, children in this study ranked the mechanoid Cozmo robot higher once they viewed a video of the robot expressing highly social emotive expressions; a feature inadequately conveyed by static images of the robot. While real-life interactions are preferable over video depictions of robots, video depictions can still provide participants with an accurate representation of different robots and allow for broader testing of different robots with larger numbers of participants. This can provide an important first step in informing future research with physically present/embodied robots. We therefore used video depictions of all robots instead of static images alone.

We sought to follow a dimensional approach, as proposed by [2], whereby teachers rated robots across a range of features, to evaluate their differentiation as a function of robot softness, and identify their influence on teachers' perceptions and expectations of the utility of different social robots in education. We specified the following preregistered hypotheses.

1) RQ1: How do teachers perceive soft and hard robots?: Based on [8] and [15], we expected soft robots to be rated as less competent (**H1.1**), but more physically (**H1.2**) and socially warm (**H1.3**) than hard robots. Second, as [16] reported higher perceived discomfort with a robot made of fur than wood (measured through ratings of awkward, scary, strange, awful, dangerous, and aggressive), we expected soft robots to be perceived as more uncomfortable to be around (**H1.4**) and creepy (**H1.5**). Following [15], we expected soft robots to be perceived as safer (**H1.6**), friendlier (**H1.7**), and more social than hard robots (**H1.8**). We did not specify directional hypotheses for human-likeness in form and motion, social competence, intelligence[1], and social intelligence.

2) RQ2: Are soft robots suitable for different educational tasks, populations, contexts, and roles than hard robots?: We expected soft robots to be perceived as more suitable for tasks that require emotional and social skills, rather than competence, as soft robots have been perceived to be lower in competence, but higher in warmth [8], [15]. Hence, we expected soft robots to be more suitable for helping someone feel better about tasks they find difficult (**H2.1**) and listening to them read (**H2.2**), but less suitable for helping them learn mathematics (**H2.3**). Co-creating an artwork with a child could fall into the category of competence or emotional and social skills, as the robot could also serve as a source of inspiration to the child.

We expected soft robots to be perceived as more suitable for children than hard robots, as their soft exteriors might resemble plush toys, which adults might associate with child-friendliness (**H2.4**). Additionally, a soft exterior might make robots seem safer, which could be a second driver behind perceiving them as more suitable for children. We expected hard robots to be perceived as more suitable for adults than soft robots (**H2.5**). We did not specify directional hypotheses regarding the suitability for individuals or groups, nor for role expectations.

3) RQ3: Do teachers favour soft or hard robots for the use in their classroom and their students' use?: Lastly, we explored the usability of the robot types by asking teachers about their own preferences and perceptions of their students' preferences for particular robots in the classroom.

II. Materials and methods

A. Participants

Data from 120 teachers were collected (83 female, 37 male, m_{age} = 42.8, SD_{age} = 11.09) from Prolific [37] (*n* = 87) and the broader community (*n* = 33). Teachers were from Europe (*n* = 113) and North America (*n* = 7, **Supplement B** in [17]). One-hundred-and-two teachers taught at a public school, 11 at a private school, four at a religious school and three at other institutions (e.g., extracurricular education providers). Teachers taught a range of subjects (e.g., Languages, Mathematics, Social and Natural sciences) and taught pupils between 3 and 18+ years of age (3-5-years: *n* = 22; 6-8-years: *n* = 33; 9-11-years: *n* = 58; 12-14-years: *n* = 54; 15-18-years: *n* = 61; above 18-years: *n* = 3). Sample size justifications can be found in our preregistration [17]. The study was approved by the local ethics committee (2024-N-12).

[1]We neglected to preregister a hypothesis for intelligence. We therefore include analyses of this factor as exploratory.

B. Design and procedure

We used a within-participants, repeated measures, mixed-methods design. Data were collected via an online Qualtrics survey [36]. Teachers were informed about the goal of the study, but were otherwise naive about our research questions. Each teacher rated 12 videos featuring six hard and six soft robots. Each video was rated across 16 dimensions (bad/good, comfortable to be around, competence, creepy/cute, familiarity to similar objects, familiarity to the robot, friendliness, human-likeness in form and motion, intelligence, physical warmth, safety, social competence, social intelligence, socialness, and social warmth). Additionally, teachers were asked to rank the robots in comparison to each other on all dimensions. The order of video and dimension presentation were counterbalanced across participants to mitigate order effects. Additionally, teachers were asked to rate the suitability of each robot for four different tasks (order of task was counterbalanced), two age groups, two contexts, and four educational roles. For each question, teachers were asked to elaborate on their choice in an open-ended question. Lastly, teachers were asked to choose three robots to use in their own classroom and three robots that their students would be most excited to use. The survey took an average of 45 minutes to complete.

C. Materials

We selected 12 robots for this study, six of them with a hard exterior (e.g., plastic) and six with a soft exterior (e.g., fur, fabric). In engineering sciences robots are defined as 'soft' if they comprise soft materials in both their general construction, and for their actuators (e.g., hoists, pneumatic). However, we classified robots as soft or hard based on the materials used to construct the robots' external shells. This decision was taken due to a lack of "classical" soft robots in the social robotics domain. We selected some of the most prominently used robots in past education research, including the NAO [18], Pepper [18], Baxter [19], Moxie [20], Miro [21], Cozmo [22], Tega [23], and Paro [24] robots, and also added some lesser known soft robots, including Keepon [25], Blossom [6], Tofu [26], and Huggable [27] (**Fig. 1**). The selected hard robots generally possessed more human-like forms and advanced physical capabilities (e.g., arms) compared to the soft robots. While this morphological variation may influence perceptions of competence and human-likeness, our selection aimed to include prominently used social robots in educational research [1], [7], [28], reflecting the current diversity in robot embodiments.

Video stimuli depicting each robot were extracted from YouTube [29], muted to avoid confounding effects from speech, and edited to 4-5 seconds and in a way to ensure minimal background context. Videos were intentionally curated to depict the robot in motion to showcase its functional capabilities and to provide ecologically valid depictions of these robots already available in the public domain (see [17] for videos). We deliberately chose to showcase the robots in a neutral context as opposed to a teaching context, as this would have influenced participants imagination of what educational tasks each robot might be suitable for.

D. Measures

1) Control variables: We assessed teachers' general attitudes towards robots, using the validated, 20-items General Attitudes Towards Robots Scale (GAToRS; [30]), which measures people's personal and societal level positive and negative attitudes towards robots on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree; see [17] for items). Additionally, we assessed teachers' prior contact with robots, using an adapted version of the items used by [31] (see [17] for items). Lastly, we recorded the number of times each video was watched and implemented two attention checks throughout the survey.

2) RQ1: Robot perception: We assessed teachers' perceptions of the robots on 16 dimensions. We derived the dimensions from previous work which identified important robot features based on survey data [8], [12], [15], [32]–[34] and theoretical considerations [2], [33]. Each dimension was measured on a slider from 0 to 100, with the end points being polar opposite attributes (e.g., bad/good) or low/high. We asked teachers how familiar they were with the robots [32], how familiar they were to objects that are similar to the robots (adapted from [2]), and how comfortable or uncomfortable they would feel around the robots [15]. Previous research examined perceptions of general competence, using items that describe robots as competent, confident, independent, competitive, and intelligent [33]. However, this conflates perceptions of competence in separate domains. For instance, a robot (or person) can be perceived as competent in completing complex tasks, but not in empathizing with others. Therefore, we separately examined general competence, social competence, intelligence, and social intelligence. Likewise, instead of assessing a general construct of warmth which in previous work was assessed through describing a robot as tolerant, warm, and good natured [8], we separately assessed physical warmth, social warmth, and a robot's perceived goodness/badness. Additionally, we measured friendliness and cuteness/creepiness (adapted from [10]), the human-likeness of form and motion, and socialness (adapted from [2]). Lastly, we asked teachers how safe they perceived each robot as being, since safety is rarely addressed in educational HRI studies, but an important construct when introducing robots to a vulnerable population such as children [7].

3) RQ2: Suitability: We asked teachers to rate the suitability of each robot for four different tasks on a slider from 0 to 100. These tasks included helping a child: (i) learn mathematics; (ii) create an artwork; (iii) feel better about tasks they find difficult; and (iv) listening to a child read. Those tasks were selected as they are common in HRI research (e.g., reading, [4]; mathematics, [35]) or reflect potential tasks a robot could be used for in education contexts (i.e., making a child feel better; helping to create art). Teachers also had the opportunity to indicate any other task the robots could be suitable for.

Next, across three tasks, teachers were asked to sort each robot into one of four suitability categories respectively, using a drag-and-drop function: (i) children, adults, both, or neither; (ii) individual use, groups, both, or neither; and (iii) teacher – who tells students about things; tutor – who guides students through tasks; peer – who works on tasks with students; or novice/tutee – who is taught things by students. The roles were extracted from prominent educational roles of robots in previous work [1], [7].

4) RQ3: Preference for robots: Lastly, we asked teachers to pick three robots that they would potentially like to use in their classroom and three robots they thought their students would be most excited to use. We included this question as an exploratory measure to capture teachers' general attitude towards both robot categories. This approach acknowledges that teachers' overall preferences for certain robots may reflect a holistic evaluation rather than a sum of individual dimensions. For example, a robot's perceived warmth may outweigh its technical competence when considering its suitability for younger learners, while the reverse may be true for adult-focused learning tasks.

E. Data analysis

1) Preliminary analysis: All analyses were conducted in R [36] using RStudio [37]. We conducted correlations between the continuous study measures (**Supplement C**). Means and standard deviations of all study measures are reported in **Supplement D**. Both attention checks were passed by all participants. Teachers watched each robot video at least once and 2.21 times on average (SD = 0.73; **Supplement A**, **Fig. S1B**). Teachers varied in their exposure to robots of different kinds, indicating that we sourced a representative sample (**Supplement B**, **S2**).

2) Main analysis: All analyses were performed using mixed-effect models using the lme4 [38] or ordinal [39] packages. Each model included a (varying) dependent outcome variable, a fixed effect for the variable robot type (using sum-to-zero contrasts; soft coded as -1, hard as 1) and – in some models – a second independent variable and their interaction. The repeated measures nature of the data was modelled by including a random slope for robot type across participants, a per-participant random intercept, and a random intercept for robot. Confidence intervals were calculated as Profile or Wald intervals.

RQ1: Perception of robots. We assessed differences between teachers' perceptions of hard and soft robots across 16 different dimensions according to three preregistered analyses: (i) Differences in the continuous ratings (0-100) of each dimension; (ii) differences in the ranking (1-12) of robots in each dimension; and (iii) differences in the likelihood of being ranked first (1) versus not being ranked first (0) in each dimension. The rating data were analysed with linear mixed effects models (LMM), using the lmer function of the lme4 package. The ranking data were analyzed using cumulative link mixed models (CLMM), using the clmm function of the ordinal package. Third, we ran generalised linear mixed effects models (GLMM) on the likelihood of being ranked first versus not being ranked first, using the glmer function of the lme4 package.

RQ2: Suitability for tasks, ages, context, and roles. We assessed differences in the perceived suitability of hard versus soft robots for different tasks, age groups, contexts, and roles. The task suitability data were analysed with four LMMs, each including participants' rating (0-100) as the dependent variable. The age, context, and role suitability data were analysed with GLMMs. The models included participants' suitability rating as the dependent variable (coded as 0 if not suitable, 1 if suitable), a fixed effect for the variable age, context, or role respectively, and their interaction with robot type. Contrasts were calculated using emmeans [40] with a FDR correction for multiple comparison.

RQ3: Preference for robots. We assessed differences in the preference for using hard and soft robots in the classroom from the perspective of teachers, as well as teachers' perceptions of their students' preferences. The data were analysed using a CLMM, with the variable rank as the dependent variable (ranks 1, 2, and 3, and non-ranked coded as 0), a fixed effect for the variable target group (teachers' self preference versus their students' assumed preference, using sum-to-zero contrasts), and their interaction.

3) Qualitative analysis: Qualitative analyses were performed in MAXQDA [41]. The qualitative data from teachers' open-ended answers were coded separately for each research question. We summarized teachers' ideas of suitable tasks per robot type. Teachers' justifications for their sorting according to age, context, and roles were coded for descriptive segments and summarized in word clouds (**Supplement K**). A total of 1205 text segments from 104 teachers were coded into 322 codes.

III. Results

A. RQ1: Perception of robots

Our analysis showed differences in the rating-based perception of hard and soft robots in competence, familiarity, human-likeness, intelligence, and physical warmth (**Fig. 2**, **Fig. S3**). Hard robots were perceived as more competent (β = 8.56, SE = 2.20, 95% CI [4.21, 12.91][2]; **H1.1**), more familiar (β = 4.46, SE = 2.22, 95% CI [0.07, 8.86]), more human-like in their form (β = 9.50, SE = 4.67, 95% CI [0.39, 18.61]), and more intelligent (β = 9.18, SE = 2.30, 95% CI [4.65, 13.71]), but less physically warm (β = -7.06, SE = 2.98, 95% CI = [-12.92, -1.21]; **H1.2**) than soft robots. No significant differences between hard and soft robots emerged in the perception of all other dimensions (see **Table S6** for model estimates).

These results were partly confirmed by the ranking-based analyses. We found hard robots being ranked higher in competence (β = -1.22, SE = 0.28, $p < .001$, 95% CI [-1.77, -0.68]; **H1.1**), human-likeness in form (β = -1.24, SE = 0.56, p = .03, 95% CI [-2.34, -0.14]), and social intelligence (β =

[2]We report 95% CI instead of p-values for the lme models, as the authors of the lme4 package chose not to include p-values due to their complex and potentially misleading calculation for these models [12].

-0.62, SE = 0.28, p = .03, 95% CI [-1.16, -0.07]), but lower in physical warmth (β = 0.74, SE = 0.29, p = .01, 95% CI [-0.17, 1.32]; **H1.2**). We did not find significant differences between hard and soft robots in all other dimensions (see **Table S6** for model estimates).

Lastly, we assessed differences between hard and soft robots in the likelihood of being ranked first for each dimension. We found hard robots more likely to be ranked first in regard to competence (β = 1.25, SE = 0.53, p = .02, 95% CI [0.20, 2.29]; **H1.1**) and human-likeness (β = 1.41, SE = 0.55, p = .01, 95% CI [0.32, 2.49]), and soft robots more likely to be ranked first regarding their cuteness (β = -0.98, SE = 0.33, p = .003, 95% CI [-1.62, -0.34]; **H1.5**). We did not find significant differences in the likelihood of being ranked first in all other dimensions (see **Table S6** for model estimates). To summarize, hard robots were evaluated as more competent, familiar, and human-like, but less physically warm than soft robots. Additionally, hard robots were rated as more intelligent and socially intelligent, but less cute than soft robots, but these differences were only present in the ranking data and not in the rating data.

B. RQ2: Suitability

1) Task suitability: Our analyses showed significant differences in the perceived task suitability of hard and soft robots for some but not all tasks (**Fig. 3A**, **Fig. S4**). Hard robots were perceived as more suitable for helping a child learn mathematics (β = 11.86, SE = 3.62, 95% CI [4.78, 18.94]; **H2.3**) or create an artwork (β = 6.02, SE = 3.01, 95% CI [0.10, 11.93]). There were no differences in the perceived suitability of hard and soft robots with regards to listening to a child read (**H2.2**) or helping a child to feel better about difficult tasks (**H2.1**; see **Table S7** for model estimates).

Qualitative results. Additional tasks that were mentioned as suitable for soft robots were primarily emotional support tasks (n = 84), such as providing encouragement or comfort, helping students to verbalize their feelings and regulate emotions on a difficult day, and providing stress release and cuddles. Additionally, teachers mentioned: (i) Non-educational tasks (n = 33), such as entertaining and playing; (ii) literacy tasks (n = 20), such as reading to children or listening to them read; and (iii) creative tasks (n = 21), such as role play, drama, and singing. Teachers also mentioned that soft robots could be especially well suited for younger children (n = 18).

When thinking of suitable tasks for hard robots, teachers primarily mentioned general tasks (n = 61), such as assessing and monitoring children, providing general help in the classroom, or recalling information. Additionally, teachers mentioned: (i) STEM related tasks (n = 59), such as helping with experiments, teaching about sciences, and coding; (ii) physical tasks (n = 39), such as leading or encouraging physical education, and (iii) teaching tasks (n = 28), such as teaching a variety of subjects, providing instructions, or functioning as an assistant to the teacher (see [17] for codebook).

2) Age suitability: We compared whether hard and soft robots were more likely to be categorized as suitable for children, adults or neither children nor adults (**Fig. 3B**, **Fig. S5**). Teachers were more likely to indicate that hard robots were suitable for adults (β = 3.40, SE = 0.24, p < .001; **H2.5**) and that soft robots were suitable for children (β = -0.47, SE = 0.19, p = .02; **H2.4**). Teachers were also more likely to indicate that soft robots were unsuitable for adults and children compared to hard robots (β = -0.82, SE = 0.22, p < .001; see **Table S8** for contrasts).

Qualitative results. Robots that were deemed suitable for children were primarily described as toy- or teddy-like (n = 27), cute (n = 24), and soft and cuddly (n = 15). Robots suitable for adults were mainly described as human-like (n = 5), functional (n = 4), and mechanical (n = 4; **Supplement K**).

3) Context suitability: We compared whether hard and soft robots were more likely to be categorized as suitable for use in individual or group settings, or neither individual nor group settings (**Fig. 3B**, **Fig. S6**). Teachers were more likely to indicate that hard robots were suitable for groups (β = 1.83, SE = 0.15, p < .001) and that soft robots were suitable for individuals (β = -0.57, SE = 0.15, p < .001). Teachers were also more likely to indicate that soft robots were unsuitable for groups and individuals compared to hard robots (β = -0.53, SE = 0.18, p = .003; see **Table S9** for contrasts).

Qualitative results. Robots that were considered as suitable for individual use were primarily described as toy-like

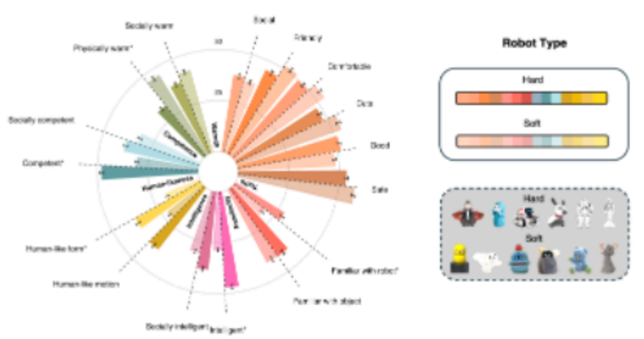

Fig. 2. Spider plot of means per dimension with standard errors by robot type. Dark shades correspond to hard robots and lighter shades correspond to soft robots. *Significant differences between hard and soft robots.

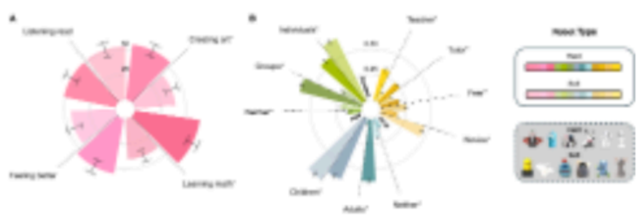

Fig. 3. Suitability ratings. **A** shows the means of dimensions with standard errors by robot type. **B** shows the proportions of suitability per age group, context, and role. Dark shades correspond to hard robots and lighter shades correspond to soft robots. *Significant differences between hard and soft robots.

(n = 12), small (n = 9), cuddly (n = 7), and cute (n = 7). Additionally, teachers mentioned the ability of these robots to form a personal bond (n = 6). Robots suitable for a group context were mainly described as having a typical robotic appearance (n = 7), being large (n = 4) and human-like (n = 4; **Supplement K**).

4) Role suitability: We compared whether hard and soft robots were more likely to be categorized as suitable for the role of a teacher, tutor, peer, and novice (**Fig. 3B**, **Fig. S7**). Teachers were more likely to categorize hard robots as suitable for the role of a teacher (β = 1.53, SE = 0.15, $p < .001$) and tutor (β = 0.86, SE = 0.14, $p < .001$) than soft robots. On the other hand, teachers were more likely to categorize soft robots as suitable for the role of a novice (β = -1.46, SE = 0.13, $p < .001$) and peer (β = -0.18, SE = 0.07, p = .01) than hard robots (see **Table S10** for contrasts).

Qualitative results. Robots that were deemed suitable for the role of a novice were mostly described as simple (i.e., basic abilities, low competence; n = 8), toy-like (n = 5), and unintelligent (n = 4). Robots that were considered suitable for the role of a peer were also described as toy-like (n = 7), but also as friendly (n = 3), cute (n = 3), and a companion (n = 3). Robots suitable for the role of a tutor were described as human-like (n = 3), mobile (n = 2), functional (n = 2), and intelligent (n = 2). Lastly, robots suitable for the role of a teacher were described as human-like (n = 19), of a typical "robotic" appearance (n = 4), and complex (n = 3; **Supplement K**).

5) RQ3: Preference for hard and soft robots: We asked teachers about their preference for the robots and to indicate what they thought their students' preferences would be (**Fig. S8**, **Fig. S9**). Teachers were more likely to prefer hard robots over soft robots for usage in their classroom (β = 1.47, SE = 0.53, p = .003) and this same preference endured for these teachers' assumed preferences of their students (β = 1.67, SE = 0.53, p = .01; see **Table S11** for contrasts).

IV. Discussion

In this study, we sought to address how soft and hard robot features influence teachers' perceptions of robot suitability for different education applications. Our results directly inform initial theoretical considerations for the application of different robot types in different education contexts. We found that hard robots were perceived as more competent, human-like, intelligent, socially intelligent, and familiar than soft robots. Soft robots on the other hand were perceived as more physically warm and cute. These differences in perception possibly drove differences in the educational suitability of these robots: Hard robots were perceived as more suitable for "technical" tasks, such as mathematics, art, and adopting a teacher/tutor role for supporting the learning of adults or groups. Soft robots were more likely to be evaluated as suitable for use with younger learners in individual learning contexts and playing the role of a co-learner/novice.

A. Differences in perception

We replicated and extended prior work [8], [15], showing that soft robots were perceived as less competent than hard robots, even when depicting the functional capabilities of these robots in videos. Additionally, teachers rated hard robots as more human-like, which is intuitive given that three of the selected hard robots were humanoid robots. Reference [11] found a robot's human-likeness to influence its perceived competence, warmth, and discomfort, indicating a relationship between a robot's human-likeness and its competence. This link is supported by our qualitative findings, as teachers deemed hard robots (who were rated as more competent) as more suitable for adults, groups, and the role of a tutor and teacher, and described robots suitable for these contexts overwhelmingly as human-like. Soft robots on the other hand, were deemed as more suitable for individual use with children, in the role of a co-learner or novice. Robots suitable for these contexts were described as toy-like, simple, and unintelligent, showcasing the other end of the competence and human-likeness spectrum. Teachers' perceptions of a lack of intelligence could in fact be a benefit in some situations, as this could contribute to a less intimidating peer/novice context, which could in turn promote further learner engagement [42].

Additionally, we found that soft robots were perceived as physically warmer but not socially warmer than hard robots. This is in line with findings by [8], [15], who did not find a difference in (general) warmth between plush and hard robots, nor between plush, wood, and plastic robots. This could indicate that both hard and soft robots can elicit warmth in different ways. While soft robots could provide physical comfort and emotional support through their softer shell, hard robots might provide comfort through their advanced interaction capabilities – including the ability to verbally express emotions. This is supported by our qualitative findings, in which teachers indicated that soft robots would be suitable for emotional-support tasks, such as providing comfort, cuddling, and helping children to vocalize their feelings, while hard robots were deemed suitable for helping in different ways, such as through organizing, physical help, or teaching different subjects.

We also found differences between hard and soft robots in intelligence, social intelligence, and cuteness. While intuitive, these differences were not robust across analyses. This could indicate that teachers changed their ratings based on the direct comparison between robots, or that teachers did not pay close enough attention to the ranking, as they already rated the robots before. While these findings need to be treated with caution, mentions of intelligence and cuteness were also present in teachers' qualitative answers. Here, cuteness and/or being unintelligent were mentioned in connection with a robot being suitable for children, individuals, and the roles of a peer and novice, while intelligence was used to describe suitable robots for groups and the role of a tutor and teacher. We found these contexts to be differently suitable for hard and soft robots, providing some support for the robustness of differences in cuteness and intelligence between hard and soft robots.

B. Differences in suitability

The differences in the perceived competence, human-likeness, and warmth between hard and soft robots seemed to be the main driver behind the differences in perceived suitability of these robots for different tasks, age groups, contexts, and roles. Hard robots were perceived as more suitable for contexts that could require more competence, such as helping a child learn mathematics or creating an artwork, interacting with adults and groups, and assuming the role of a teacher and tutor. Such competence appeared to be demonstrated by robot size, interaction capabilities, or movement, as teachers mentioned these features but also general competence as criteria. One teacher explained that "*robots that resemble humans and look like they are more intelligent*" were more suitable to adopt a teacher/tutor role (29f, public school, UK).

Soft robots were perceived as more suitable for contexts that possibly require more emotional skills, such as interacting with individuals, children, and taking the role of a co-learner or novice. This is supported by teachers' qualitative answers: "*I'm basing these [rankings] on children's needs for [robots] that have physical warmth and that older kids like something that is a little more 'grown up' or may be capable of more complex tasks*" (33f, additional needs education, Ireland). Another teacher explained that "*the really soft and fuzzy ones seem like it would be easier for one person [to] develop kind of a personal relationship with them if they weren't in a group*" (76m, public school, USA).

Regardless of the perceived differences between hard and soft robots and the resulting differences in perceived suitability, teachers indicated a general preference for hard robots for use in their classroom and this preference did not change, regardless of whether teachers chose for themselves or for their students. The preference for hard robots did not depend on the subject taught or pupils' age. This could indicate that teachers may lack an understanding of how robots can be used or that they might not have clear ideas about how to use them across different learning contexts. This general preference for hard robots is in line with the perceived better suitability of these robots for group contexts and is thus rather intuitive. However, if pupils indeed prefer hard robots over soft robots – as expected by their teachers – must be empirically established by future work.

C. Limitations

There are several limitations and confounding variables to consider. First, we did not include a humanoid soft robot in our sample, such as Sophia [43]. This decision was made deliberately since these robots are rather uncommon and usually trigger an uncanny valley response (i.e., are evaluated as eerie; [44]). Consequently, most of the soft robots used in our study were toy-like robots, sometimes closely resembling plush toys. While this appearance could have driven the differences in perception and suitability, it remains open how more technological- or human-looking soft robots, made from silicon, rubber, or "skin" would be situated in education contexts. However, as mentioned above, perceptions of reduced competence among soft education robots could be beneficial in some contexts. Second, while the diverse technical, communicational, and morphic properties of the selected robots add ecological validity, they also introduce confounding factors like size, shape, and communication abilities [12]. Despite 'low-tech' hard robots being uncommon commercially, future research comparing functionally similar hard and soft robots would help to isolate complexity effects. Nevertheless, the broad sampling in this study helped mitigate confounding effects. Third, we used video depictions of the robots rather than real-life interactions. While this allowed us to examine participants' perceptions of a broader range of robots, this decision also limits the ecological validity as people might be differently affected by physically embodied interactions compared to videos. Our findings should be replicated in an in-person experiment. Lastly, we assessed students' preferences for hard or soft robots through the perspective of their teachers. Future work should assess whether students' preference aligns with their teachers' expectations.

D. Conclusion and Future Work

Hard and soft robots differed in their perceived competence, human-likeness, and physical warmth. These differences seemed to shape for which educational tasks, age groups, contexts, and roles these robots are deemed suitable for. Our findings indicate that different types of robots could be deployed for different educational contexts and that not every deployment requires a highly complex and costly hard robot. This opens the possibility of using cheaper and less complex soft robots in certain education contexts, thus making education robots more accessible and scalable. However, teachers generally preferred hard robots over soft robots for use in their classrooms, highlighting challenges in the broader acceptance of soft robots as a useful education aid. A possible solution could be "hybrid" robots that combine the competence and functionality of hard robots with the higher physical warmth and emotional support capabilities of soft robots. Additionally, we should promote robot literacy to help teachers understand the possibilities that different robots provide and what current evidence tells us about their effectiveness in different contexts. This could increase the acceptance of different robot types and help teachers understand that even an incompetent novice robot could be an effective educational tool.

Our findings contribute to the literature in multiple ways. First, from a theoretical perspective, we provided new evidence on how robot softness affects perceptions of competence and warmth, which are key dimensions in educational contexts. While prior research identified these differences, we extended this knowledge by examining them specifically in an educational setting using video depictions of robots. Second, from a methodological perspective, we demonstrated that a two-dimensional framework focusing solely on competence and warmth fails to capture critical variability across robot types. Instead, a multi-dimensional approach, incorporating factors such as human-likeness, intelligence, and physical warmth, provides a more comprehensive understanding of these perceptions and can guide future research on robot design and evaluation. Future work should build on these findings by investigating how these perceptions translate into practical, in-person classroom settings.

ACKNOWLEDGMENT

We would like to thank Dr. Ryssa Moffat. **Funding:** Funding was received by the Social Brain Sciences Lab. **Open Access:** The authors have applied a Creative Commons Attribution (CC BY) licence. **Data access:** Data, including the supplementary material are available on the OSF [17]. **Preregistration:** Our hypotheses, methods, and data analyses were preregistered [17].

REFERENCES

[1] T. Belpaeme, J. Kennedy, A. Ramachandran, B. Scassellati, and F. Tanaka, "Social robots for education: A review," *Science Robotics*, vol. 3, no. 21, 2018.

[2] E. S. Cross and R. Ramsey, "Mind meets machine: Towards a cognitive science of human–machine interactions," *Trends in Cognitive Sciences*, vol. 25, no. 3, pp. 200–212, March 2021.

[3] K. J. Rohlfing, N. Altvater-Mackensen, N. Caruana, R. Van Den Berghe, B. Bruno, N. F. Tolksdorf, and A. Hanulíková, "Social/dialogical roles of social robots in supporting children's learning of language and literacy—a review and analysis of innovative roles," *Frontiers in Robotics and AI*, vol. 9, p. 971749, Oct. 2022.

[4] N. Caruana, R. Moffat, A. Miguel-Blanco, and E. S. Cross, "Perceptions of intelligence & sentience shape children's interactions with robot reading companions," *Scientific Reports*, vol. 13, no. 1, p. 7341, May 2023.

[5] F. Stella and J. Hughes, "The science of soft robot design: A review of motivations, methods and enabling technologies," *Frontiers in Robotics and AI*, vol. 9, p. 1059026, Jan. 2023.

[6] M. Suguitan and G. Hoffman, "Blossom: A handcrafted open-source robot," *ACM Transactions on Human-Robot Interaction*, vol. 8, no. 1, pp. 1–27, Mar. 2019.

[7] H. Woo, G. K. LeTendre, T. Pham-Shouse, and Y. Xiong, "The use of social robots in classrooms: A review of field-based studies," *Educational Research Review*, vol. 33, p. 100388, Jun. 2021.

[8] B. Reeves, J. Hancock, and X. Liu, "Social robots are like real people: First impressions, attributes, and stereotyping of social robots," *Technology, Mind, and Behavior*, vol. 1, no. 1, pp. 1–11, Oct. 2020.

[9] A. M. Harris-Watson, L. E. Larson, N. Lauharatanahirun, L. A. DeChurch, and N. S. Contractor, "Social perception in human-ai teams: Warmth and competence predict receptivity to ai teammates," *Computers in Human Behavior*, vol. 145, p. 107765, August 2023.

[10] E. Liberman-Pincu, E. D. Van Grondelle, and T. Oron-Gilad, "Designing robots with the context in mind- one design does not fit all," in *Human-friendly robotics 2022*, P. Borja, C. Della Santina, L. Peternel, and E. Torta, Eds. Springer International Publishing, 2023, pp. 105–119.

[11] L. Kunold, N. Bock, and A. Rosenthal-von Der Pütten, "Not all robots are evaluated equally: The impact of morphological features on robots' assessment through capability attributions," *ACM Transactions on Human-Robot Interaction*, vol. 12, no. 1, pp. 1–31, March 2023.

[12] E. Liberman-Pincu, Y. Parmet, and T. Oron-Gilad, "Judging a socially assistive robot by its cover: The effect of body structure, outline, and color on users' perception," *ACM Transactions on Human-Robot Interaction*, vol. 12, no. 2, pp. 1–26, June 2023.

[13] V. Venkatesh, J. Y. L. Thong, and X. Xu, "Unified theory of acceptance and use of technology: A synthesis and the road ahead," *Journal of the Association for Information Systems*, vol. 17, no. 5, pp. 328–376, May 2016.

[14] J. M. Beer, A. Prakash, T. L. Mitzner, and W. A. Rogers, "Understanding robot acceptance," Georgia Institute of Technology, Tech. Rep., 2011.

[15] S. Van Waveren, L. Björklund, E. J. Carter, and I. Leite, "Knock on wood: The effects of material choice on the perception of social robots," in *Social Robotics*, M. Salichs, S. Ge, E. Barakova, J.-J. Cabibihan, A. Wagner, Castro-González, and H. He, Eds. Springer International Publishing, 2019, pp. 211–221.

[16] L. Björklund, "Knock on wood does material choice change the social perception of robots?" KTH Royal Institute of Technology, Tech. Rep., 2018.

[17] L. M. Leisten, N. Caruana, and E. S. Cross, "TRESH - Teachers' Reactions to Soft and Hard Robots in Education," January 2024. [Online]. Available: https://doi.org/10.17605/OSF.IO/K8GT4

[18] "Aldebaran robotics," 2023, https://www.aldebaran.com/en/pepper.

[19] "Rethink robotics," Accessed: 2024-09-17, 2024, https://rethinkrobotics.com/.

[20] "Moxie robot," 2024, https://moxierobot.com/.

[21] "Miro-e," 2024, https://www.miro-e.com/.

[22] "Anki," Accessed: 2023-10-30, 2023, https://ankicozmorobot.com/.

[23] "Tega: A new social robot platform," Accessed: 2024-09-17, 2024, https://www.media.mit.edu/projects/tega-a-new-robot-platform-for-long-term-interaction/overview/.

[24] "Aist," 2024, https://www.aist.go.jp/index_en.html.

[25] "Beatbots," 2024, https://beatbots.net/keepon-pro.

[26] "Tofu – personal robots group," Accessed: 2024-09-17, 2024, https://robots.media.mit.edu/portfolio/tofu/.

[27] M. P. R. Group, "Huggable – personal robots group," 2024.

[28] H. Mahdi, S. A. Akgun, S. Saleh, and K. Dautenhahn, "A survey on the design and evolution of social robots — past, present and future," *Robotics and Autonomous Systems*, vol. 156, p. 104193, October 2022.

[29] "Youtube," 2024, https://www.youtube.com/.

[30] M. Koverola, A. Kunnari, J. Sundvall, and M. Laakasuo, "General attitudes towards robots scale (gators): A new instrument for social surveys," *International Journal of Social Robotics*, vol. 14, no. 7, pp. 1559–1581, Sep. 2022.

[31] A. C. Horstmann and N. C. Krämer, "Great expectations? relation of previous experiences with social robots in real life or in the media and expectancies based on qualitative and quantitative assessment," *Frontiers in Psychology*, vol. 10, p. 939, April 2019.

[32] R. Baddoura and G. Venture, "Social vs. useful hri: Experiencing the familiar, perceiving the robot as a sociable partner and responding to its actions," *International Journal of Social Robotics*, vol. 5, no. 4, pp. 529–547, November 2013.

[33] S. T. Fiske, A. J. Cuddy, and P. Glick, "Universal dimensions of social cognition: Warmth and competence," *Trends in Cognitive Sciences*, vol. 11, no. 2, pp. 77–83, February 2007.

[34] K. S. Haring, D. Silvera-Tawil, T. Takahashi, K. Watanabe, and M. Velonaki, "How people perceive different robot types: A direct comparison of an android, humanoid, and non-biomimetic robot," in *2016 8th International Conference on Knowledge and Smart Technology (KST)*, February 2016, pp. 265–270.

[35] E. Lopez-Caudana, P. Ponce, L. Cervera, S. Iza, and N. Mazon, "Robotic platform for teaching maths in junior high school," *International Journal on Interactive Design and Manufacturing (IJIDeM)*, vol. 12, no. 4, pp. 1349–1360, Nov. 2018.

[36] "R core team," 2024, https://www.r-project.org/.

[37] "Rstudio," 2024, https://www.rstudio.com/.

[38] D. Bates, M. Maechler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[39] R. Christensen, "Ordinal-regression models for ordinal data," 2023.

[40] R. Lenth, "emmeans: Estimated marginal means, aka least-squares means," 2024.

[41] MAXQDA, "Maxqda 2024," 2024.

[42] C. Zaga, M. Lohse, K. P. Truong, and V. Evers, "The effect of a robot's social character on children's task engagement: Peer versus tutor," in *Social Robotics*, A. Tapus, E. André, J.-C. Martin, F. Ferland, and M. Ammi, Eds. Springer International Publishing, 2015, pp. 704–713.

[43] Hanson Robotics, "Sophia," https://www.hansonrobotics.com/sophia/, 2024, accessed: 2025-01-06.

[44] D. Ratajczyk, "Shape of the uncanny valley and emotional attitudes toward robots assessed by an analysis of youtube comments," *International Journal of Social Robotics*, vol. 14, no. 8, pp. 1787–1803, October 2022.