



# The Effect of Design Features on Relationship Quality with Embodied Conversational Agents: A Systematic Review

Kate Loveys<sup>1</sup> · Gabrielle Sebaratnam<sup>1</sup> · Mark Sagar<sup>2</sup> · Elizabeth Broadbent<sup>1</sup>

Accepted: 11 July 2020 / Published online: 24 September 2020  
© Springer Nature B.V. 2020

## Abstract

Embodied conversational agents (ECAs) are increasingly used in healthcare and other settings to improve self-management and provide companionship. Their ability to form close relationships with people is important for enhancing effectiveness and engagement. Several studies have looked at enhancing relationships with ECAs through design features focused on behaviours, appearance, or language. However, this evidence is yet to be systematically synthesized. This systematic review evaluates the effect of different design features on relationship quality with ECAs. A systematic search was conducted on electronic databases EMBASE, PsychInfo, PubMed, MEDLINE, Cochrane Library, SCOPUS, and Web of Science in January–February 2019. 43 studies were included for review that evaluated the effect of a design feature on relationship quality and social perceptions or behaviours towards an ECA. Results synthesize effective design features and lay a scientific framework for improving relationships with ECAs in healthcare and other applications. Risk of bias for included studies was generally low, however there were some limitations in the research quality pertaining to outcome measurement and the reporting of statistics. Further research is needed to understand how to make ECAs effective and engaging for all consumers.

**Keywords** Human–computer interaction · Embodied conversational agents · Relationship quality · Design features · Healthcare

## 1 Introduction

An embodied conversational agent (ECA) is a computer-based dialogue system with a virtual embodiment (full body or face-only) that typically interacts with people using multimodal communication cues (e.g. speech, text, animated facial expressions or gestures) [1]. ECAs are increasingly used across a range of industries including healthcare, education, banking, and retail. This is made possible due to

improvements over the last decade in computer processing power, computing techniques, data availability, storage, and security. ECAs show promise for improving the supply and quality of support services across a range of industries, as ECAs are scalable, inexpensive (in comparison to robots), customizable to user needs, portable to use in many environments, and available 24/7 for support.

There are several areas where ECAs are particularly well-suited to supporting people; one of which is education. ECAs are beginning to be used in education as part of online or in-person courses. For example, ECAs have been applied to teach computer programming [2], mathematics [3], literacy [4], medical diagnosis skills (as virtual patients for medical students) [5], and social communication skills to children with autism [6]. It is important that ECAs form quality relationships with students because student–teacher relationship quality affects learning engagement and academic achievement [7]. Teacher empathy and warmth are also strongly associated with learning outcomes [8]. This suggests that if ECAs are to be effective teachers, they must be able to develop positive relationships with students, and demonstrate empathy and warmth.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s12369-020-00680-7>) contains supplementary material, which is available to authorized users.

✉ Kate Loveys  
k.loveys@auckland.ac.nz  
Elizabeth Broadbent  
e.broadbent@auckland.ac.nz

<sup>1</sup> Department of Psychological Medicine, the University of Auckland School of Medicine, Building 507, Level 3, 22-30 Park Avenue, Grafton, Auckland 1023, New Zealand

<sup>2</sup> Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand

ECAs also show potential for use in commercial settings. ECAs have been used to assist with customer service tasks in retail [9], banking [10], and real estate [11] settings. This includes helping people to find information [12], make decisions about products or services [13], or solve common problems [14]. ECAs that are able to build trust and a sense of warmth with customers have been shown to improve online purchase intention [13], satisfaction with the purchase experience and company loyalty [15].

Another promising area for ECAs to provide support is in healthcare. ECAs have been shown to improve self-management in several healthcare contexts including: stress management [16], mental health [17], medication adherence [18], breastfeeding support [19], and diet and exercise programs [20, 21], including programs for overweight adults [22]. ECAs may also provide emotional, instrumental, or informational support to people, which can influence health behaviours, such as adherence to medications or exercise regimes [23, 24]. Support provided by ECAs could also help to reduce distress and subsequently reduce the activation of the sympathetic nervous system [25], which has positive implications for physiology [26], immune function [27], and stress hormones [28].

ECAs have also been developed to provide companionship to reduce loneliness, which is a risk factor for a range of poor health outcomes [29, 30]. Companion ECAs have been used by older adults living alone [31, 32], as well as hospitalized adults [33] and children [34]. Quality companionship can provide a range of psychosocial and health benefits to people (e.g. greater mental well-being, improved cardiovascular, neuroendocrine, and immune function) [35, 36], and these effects are also seen when the companion is artificial [37–39].

Quality relationships are especially important in healthcare. Similar to patient-provider relationships, an ECA's ability to form quality relationships with patients is important for enhancing intervention effectiveness and engagement. The quality of patient provider relationships has been shown to directly and indirectly affect health outcomes [40]. This includes objective disease markers such as symptom recovery [41], blood pressure, blood sugar, and functional status [42], as well as coping [43], knowledge of the illness or treatment [44], and treatment adherence [45]. A good patient-provider relationship has also been shown to improve engagement in healthcare [46]. It is therefore possible that relationships with ECAs could have similar impacts on patient outcomes. Given the relative novelty of ECAs and their application to healthcare, further research is needed to understand how to build quality relationships with ECAs in healthcare contexts, as well as the effects of these relationships.

One way to improve relationships with ECAs in healthcare contexts could be to apply principles from

doctor-patient communication. This has been proposed in a new model on robot-patient communication [47]. Effective physician communication skills include relationship building, shared decision making, and information sharing. In this model, it is suggested that background variables related to the user (demographics, health status, personality, needs, experience and abilities) and the artificial agent (appearance, voice, gender, personality, and other design cues) influence the content of the interaction (relationship building, verbal and nonverbal cues, affective and instrumental communication), which then affects patient outcomes (engagement, satisfaction, understanding, compliance, health status). This model was developed for socially assistive robots; however, the paper proposes that this model could be applied to computer-based agents as well.

Several studies have looked at design features to enhance relationships with ECAs in healthcare and other settings. Some design features are static, such as appearance, while others are more dynamic such as behaviour models. Most design features, especially detailed behavioural models, are the result of a substantial development effort. This involves researching human characteristics or behaviour, determining modeling approaches and datasets, and user testing. There can be variation in how features appear based on the decisions made during the development process. For example, empathetic facial expression may appear different between research groups because the developers used unique modeling approaches. Research has focused on behaviours [48, 49], emotional expression [50, 51], language [52, 53], personality [2], appearance [54, 55], embodiment [32, 56], and the virtual environment [50, 57]. Some studies have found that responses to design features can vary by user characteristics such as gender [57], personality [2, 58], emotional state [53], and technical abilities [59].

The effects of ECA design features on relationships can be assessed using many outcomes. For example, relationship quality is an umbrella term that refers to positive or negative feelings about a relationship [60] and incorporates related constructs (e.g. intimacy, nurturance)[61]. Relationship quality is a term applicable to different types of relationships including professional and personal. Related constructs that have been studied across the human-computer interaction (HCI) literature, include intimacy [50], social closeness [62], and therapeutic alliance [32]. Other HCI papers have also assessed rapport, which describes a relationship quality involving positive emotions, mutual attentiveness, and coordination during interactions [63].

Research has also studied the effect of design features on social perceptions and behaviours that may form part of relationship quality. Social perceptions refer to judgements of the intentions and psychological dispositions of others [64]. Some social perceptions that are related to relationship quality and have shown to be affected by design features include:

perceived trustworthiness [33, 65], warmth [54, 66], and caring [67, 68]. Design features may also influence social behaviours related to relationship quality including engagement (the degree of user involvement and interaction [69, 70])[71, 72], degree of self-disclosure [73, 74], and desire to interact again [72, 75].

The research conducted to date on how design features impact relationship quality, social perceptions and behaviours may inform a framework for improving relationships with ECAs in healthcare and other applications. However, this evidence is yet to be systematically synthesized.

## 1.1 Aim

This systematic review aimed to evaluate the effect of different design features on relationship quality and related outcomes with ECAs. This review covers research investigating the impact of design features on relationships with ECAs across a range of settings. Results will synthesize effective design features and present a scientific framework to improve relationships with ECAs in healthcare and other applications. The review will investigate the following research questions:

- (i) What design features are shown to improve relationship quality with embodied conversational agents?
- (ii) What design features are shown to improve social perceptions and behaviours towards embodied conversational agents?

## 2 Methods

### 2.1 Eligibility Criteria

For the purpose of this review, an embodied conversational agent was defined as a dialogue agent with a virtual embodiment (full body or face-only) [1]. Studies eligible for the review were required to be (1) experiments, pilot studies, or randomized controlled trials (RCTs) with a within- or between-subjects design, where (2) the population was adults aged 18 years or older from the general public or clinical populations, and (3) the intervention of interest was an embodied conversational agent used in any context, (4) comparators were an artificial agent (a robot, embodied conversational agent, or chatbot) of an alternative design, (5) the outcome of interest was relationship quality (or similar, such as social closeness, rapport, or therapeutic alliance) or social perceptions and behaviours that could affect relationship quality (e.g. perceived supportiveness, warmth, trustworthiness, self-disclosure), (6) outcomes were assessed at least once following an

interaction with the embodied conversational agent, and (7) studies were peer-reviewed journal publications or refereed conference papers.

Studies were not excluded based on the year of publication, the setting in which the intervention was delivered, or on methodological quality. Excluded studies were those presented in abstracts only, theses or dissertations (as these are not peer-reviewed publications and therefore their ability to meet publishable standards is not demonstrated), papers published in languages other than the English language or those that focused on avatars (as a virtual representation of the user), simulation games (The Sims, Second Life), animal agents, and virtual or augmented reality.

### 2.2 Search Strategy

A systematic search was conducted on electronic databases from the health and computer sciences including EMBASE, PsychInfo, PubMed, MEDLINE, Cochrane Library, SCOPUS, and Web of Science between 21 January–03 February 2019. Manual searches were completed from reference lists, citing articles, and author citations to identify additional studies. A literature search strategy was developed from topic keywords, synonyms, and test searches to identify additional terms from titles, abstracts, and subject descriptors. A subject librarian assisted with developing the search strategy. The search strategy for EMBASE is included in “Online Appendix 1”. Searches were not limited by dates as ECAs are a relatively new technology.

### 2.3 Study Selection

Search results were imported to Covidence, an online review management software developed by the Cochrane Collaboration [76]. Duplicate articles were recorded and removed. Two independent reviewers screened the title and abstract of search results against eligibility criteria (see Sect. 2.1). Next, the reviewers examined the full text articles of studies marked as eligible for inclusion and studies where eligibility was unclear. The full text articles were collected directly from academic databases. Interrater reliability was calculated by the researchers from data provided by Covidence and indicated the extent to which the reviewers matched on judgements to include or exclude papers. Interrater reliability was 84%. Disagreements about study eligibility were recorded and resolved through discussion with a third independent reviewer. Reviewers recorded the number of studies excluded at each step and reasons for exclusion. Studies determined as eligible for review were included as part of data synthesis.

## 2.4 Data Collection

Data were extracted on study design, population, intervention, comparators, outcomes, setting, timing of measurements, publication status, and results in August 2019. Authors were not contacted for missing data or to enquire about unpublished results.

## 2.5 Risk of Bias Evaluation

A risk of bias evaluation was conducted to evaluate the internal validity of included studies. Two independent reviewers assessed risk of bias in individual studies using the Revised Cochrane Risk-of-Bias Tool for Randomized Trials (RoB 2) uploaded to Covidence for online data entry. RoB 2 examines bias across five domains which cover the randomization process, deviations from intended interventions, missing outcome data, outcome measurement, and selection of reported results. An additional domain was added to evaluate outcome measure reliability and validity. For each domain,

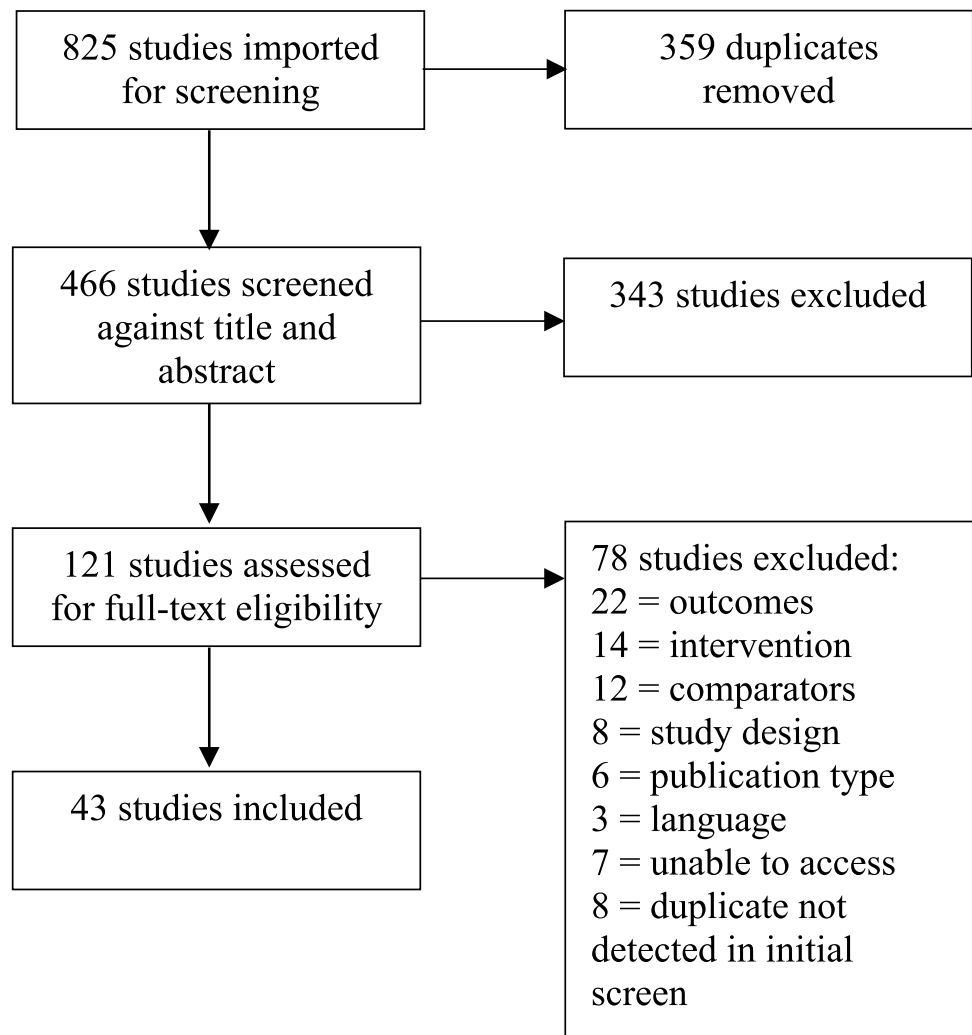
reviewers answered signaling questions that evaluated possible areas of bias, made domain-level judgements about bias risk, and predicted the direction of bias. Judgements could be ‘low’, ‘some concerns’, or ‘high’ risk of bias. After evaluating possible bias across the six domains per study, reviewers made a judgement on the overall risk of bias in the study and predicted the direction of bias.

## 3 Results

### 3.1 Study Selection

Results of the eligibility screen are depicted in Fig. 1. Searches yielded a total of 825 results. Following removal of duplicates, and an abstract only and full text eligibility screen, a total of 43 studies were included for review. Reasons for exclusion were that the study did not meet eligibility criteria due to out of scope outcomes (n=22), intervention (n=14), comparators (n=12), study design (n=8),

**Fig. 1** Flow chart of study selection (see Sects. 2.1 and 3.1 for further detail on reasons for exclusion)



publication type ( $n=6$ ), language ( $n=3$ ), it was not possible to access the paper ( $n=7$ ), or the study was a duplicate not detected in the initial screen ( $n=8$ ). Some examples of excluded studies were a study about computational modeling of human lungs for use in virtual clinical trials, a study on the gesture recognition system of an Arabic sign language program, and a study looking at predicting drug-induced arrhythmia risk from *in silico* models that use cardiac electrophysiology data.

### 3.2 Study Characteristics

Four studies contained more than one experiment from which data were extracted, making the total number of analysed studies 47. Studies ranged in sample size from 8 to 1607 participants with a median size of 57 participants (interquartile range = 33–81). Included studies were predominantly experiments ( $n=45$ ; 96%), with a between-subjects design ( $n=27$ ; 60%), that recruited adult participants ( $n=24$ ; 51%), from the general public ( $n=24$ ; 51%) or who were university students and staff ( $n=20$ ; 43%). Several studies focused on young adult ( $n=19$ ; 40%) or older adult ( $n=4$ ; 9%) populations. Most studies recruited a mixed gender sample ( $n=33$ ; 70%); however, one study recruited only female participants, and one study recruited only male participants. 12 studies ( $n=26\%$ ) did not report on the gender of the sample. Outcome assessments were mostly conducted directly after the interaction ( $n=39$ ; 83%), however some assessments were made longitudinally ( $n=5$ ; 11%), at first impression and post interaction ( $n=1$ ; 2%), after interacting for 30 days ( $n=1$ ; 2%), and longitudinally plus at two weeks follow up ( $n=1$ ; 2%). Included studies were peer-reviewed journal papers ( $n=20$ ; 47%) and conference papers ( $n=23$ ; 53%) published between 2001 and 2019. Full details of the study characteristics and results are included in “Online Appendix 2”.

### 3.3 Embodied Conversational Agent Characteristics

72 embodied conversational agents were presented across the included studies. In terms of characteristics, ECAs were most often humanlike (92%), female (53%), adult (74%), and white-‘skinned’ (67%). Depictions varied from animations to photos of a real human where facial expressions varied by text content. Characteristics of the ECAs are presented in Table 1. Examples of ECAs from some included studies are presented in Fig. 2.

### 3.4 Setting

ECAs were tested in a range of contexts including companionship ( $n=12$ ), healthcare ( $n=9$ ), commercial ( $n=9$ ), education ( $n=8$ ), gaming ( $n=5$ ), museum or conference

**Table 1** Embodied conversational agent characteristics from the 47 included studies

ECA characteristics	n	%
Appearance		
Humanlike	66	92
Robotlike	6	8
Gender		
Male	33	46
Female	38	53
Androgenous	1	1
Age		
Child	1	1
Young adult	12	17
Adult	53	74
Older adult	2	3
Unable to tell (ageless robot)	4	5
Ethnicity		
White-skinned	48	67
Tan-skinned	3	4
Dark-skinned	6	7
Hispanic	4	5
Asian	2	3
Ethnicity matched to participants (options not reported)	2	3
If robotlike, colour		
White	2	3
White–grey	3	4
Green	1	1
Other		
Appearance not reported	2	3

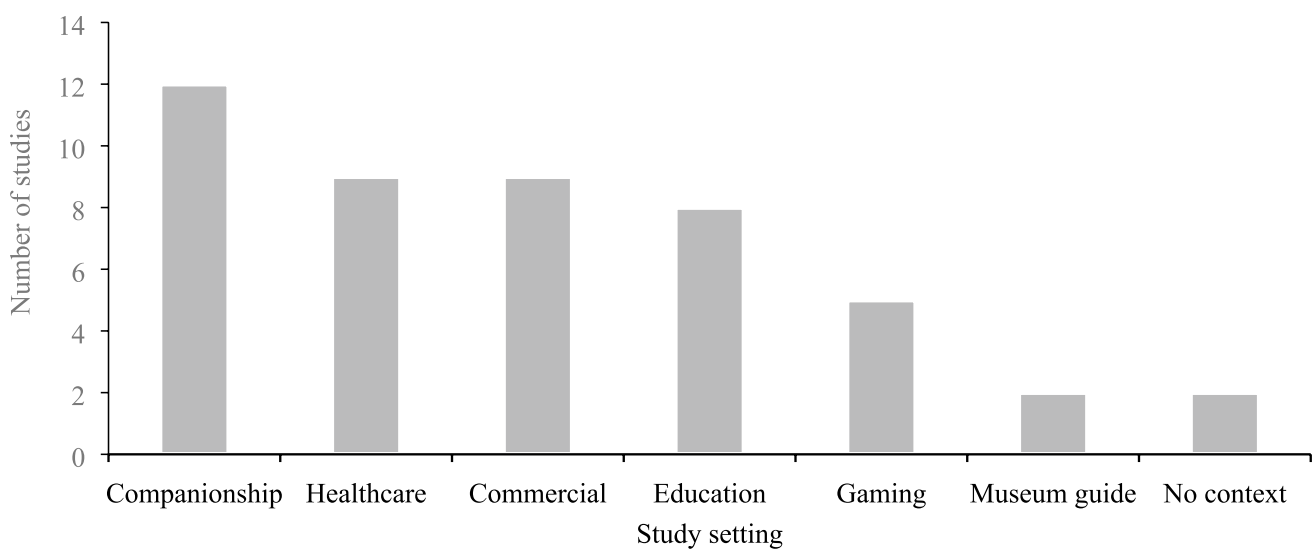
guidance ( $n=2$ ), and no context ( $n=2$ ). In healthcare settings, ECAs were used to provide counselling for mental health ( $n=4$ ) or substance misuse ( $n=1$ ), and coaching for improving exercise ( $n=3$ ) or diet ( $n=1$ ) (see Fig. 3).

### 3.5 Design Features

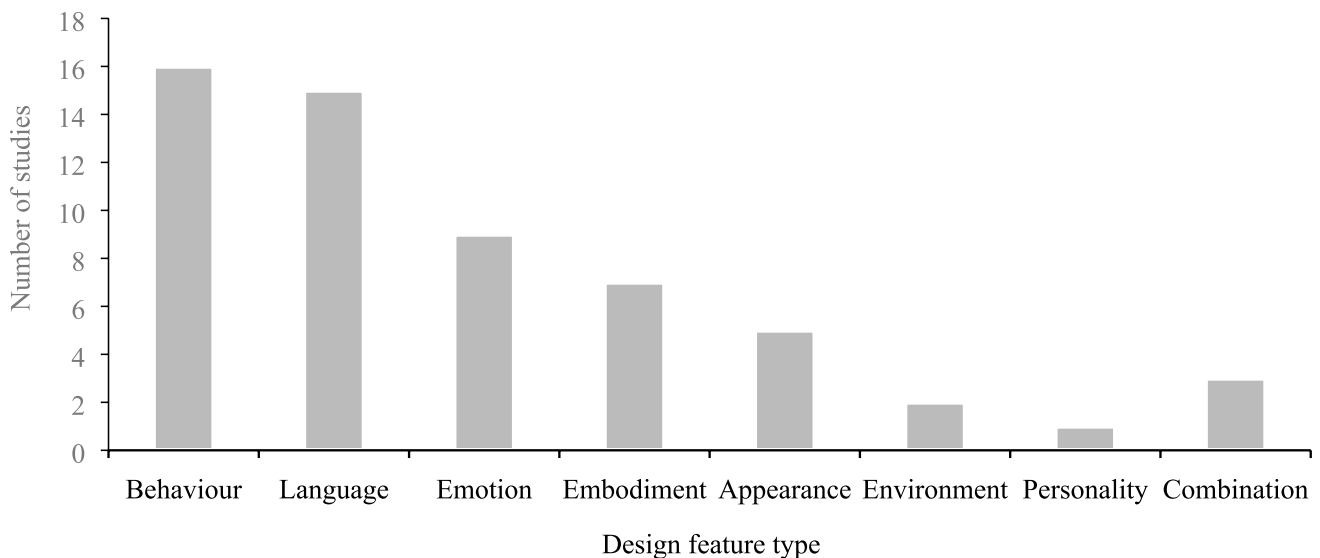
A total of 42 unique design features were evaluated across the included studies. Studies tested design features related to behaviour ( $n=16$ ), language ( $n=15$ ), emotional expression ( $n=9$ ), embodiment ( $n=7$ ), appearance ( $n=5$ ), environment ( $n=2$ ), personality ( $n=1$ ), and a combination of language and behaviour features ( $n=3$ ) (see Fig. 4). Several studies evaluated multiple design features. It was not possible to conduct a meta-analysis for relationship quality or social perceptions and behaviours as outcome measures were not adequately homogenous and reporting of results were not consistent across studies.



**Fig. 2** Examples of ECAs from included studies: **a** and **b** Rapport agents (University of Southern California Institute for Creative Technologies) [81]; **c** Laura, FitTrack Exercise Advisor (MIT) [75]; **d** ECA in likeness of performance artist Stelarc (Western Sydney University) [71]



**Fig. 3** Number of studies looking at relationships with ECAs across different settings



**Fig. 4** Number of studies evaluating the effect of design feature types on relationships with ECAs (across relationship quality, social perceptions and behaviours)

### 3.5.1 Effect of Design Features on Relationship Quality

This section presents the results for question 1 (“What design features are shown to improve relationship quality with ECAs?”). Several studies looked at the effect of design features on relationship quality ( $n = 1$ ) or similar outcomes including rapport ( $n = 16$ ), closeness ( $n = 5$ ), working alliance ( $n = 3$ ), or intimacy ( $n = 1$ ). 21 unique design features were tested that pertained to behaviour ( $n = 7$ ), language ( $n = 7$ ), emotional expression ( $n = 3$ ), the virtual environment ( $n = 2$ ), embodiment ( $n = 1$ ), and a combination of language and behaviour features ( $n = 1$ ). Results are outlined in Table 2 and described in more detail below.

### 3.5.2 Language

Seven studies investigated the effect of language features on rapport and closeness with an ECA. Language features that were shown to improve rapport included humour [49], a first-person storytelling perspective [77], ‘social reasoning’ language (including self-disclosure, acknowledgement, praise, reference to shared experiences, adherence to or violation of social norms by degree of closeness detected, questions to elicit self-disclosure) [52], and self-disclosure of humanlike stories, particularly for users with high social anxiety [53]. Context awareness improved social closeness with an ECA exercise coach, which involved demonstrating knowledge of the user’s physical activity each day [33]. There were no significant effects of user identification and high self-disclosure on relationship quality [72, 74].

### 3.5.3 Behaviour

Only 2 of the 7 behaviour features evaluated in the literature were shown to improve rapport, which were affiliative eye gaze [48] and cooperative behaviour [49]. Affiliative eye gaze referred to eye gaze that was focused predominantly on the user, as opposed to a referent object on screen. Affiliative eye gaze resulted in higher rapport ratings comparative to referential eye gaze [48]. Cooperative behaviour involved working with the user to achieve a common goal during a prisoner’s dilemma game. This was shown to build rapport better than selfish behaviour during game play [49]. Only one behavioural feature was found to have a negative effect on rapport ratings, which was high gesture amplitude (in comparison to low amplitude) [78]. There were no significant effects of mimicry of user facial expressions [79], repeated interactions [80], socially responsive behavioural feedback (head nods, smiles) [81], or high behavioural realism (e.g. breathing, blinking, head nods, posture shifts) [74, 82] on rapport.

### 3.5.4 Emotion

Emotional expression was shown to affect rapport and intimacy across four studies. Rapport ratings were higher for ECAs that demonstrated happy facial expression in comparison to sad facial expression in two experimental studies [51, 58]. One study found that positive and negative emotion in language increased ratings of intimacy compared to no emotion in language [50]. While these studies suggest that emotional expression is generally better for relationship quality with an ECA than no emotion, one study found

**Table 2** The effect of design features on relationships with ECAs (+ = increase in outcome; – = decrease in outcome; o = no effect on the relationship quality outcome,  $p < .05$ )

Design features	<i>N</i>	Subcomponents	Rapport	Closeness	Intimacy	Working alliance	Relationship quality
Language	8	Context awareness in language ( <i>vs. no context awareness</i> ) [33]		+			
	29	User identification ( <i>vs. no user identification</i> ) [72]		o			
	60	First-person storytelling perspective ( <i>vs. third-person storytelling perspective</i> ) [77]	+				
	171	High self-disclosure ( <i>vs. low self-disclosure vs. no self-disclosure</i> ) [74]	o				
	40	Self-disclosure of humanlike stories ( <i>vs. self-disclosure of computerlike stories</i> ) [53]	+				
	80	Humour ( <i>vs. no humour</i> ) [49]	+				
	228	Social reasoning language ( <i>vs. no social reasoning language</i> ) [52]	+				
Behaviour	108	Mimicry of user facial expression with contingency ( <i>vs. mimicry without contingency vs. opposite expression with contingency vs. opposite expression without contingency</i> ) [79]	o				
	171	High behavioural realism (e.g. breathing, blinking, head nods, posture shifts) ( <i>vs. low behavioural realism</i> ) [74]	o				
	83	High behavioural realism (e.g. breathing, blinking, head nods, posture shifts) ( <i>vs. low behavioural realism</i> ) [83]	o				
	79	Socially responsive behavioural feedback (e.g. head nods, smiles) ( <i>vs. no socially responsive behavioural feedback</i> ) [81]	o				
	80	Cooperative behaviour ( <i>vs. selfish behaviour</i> ) [49]	+				
	20	Repeated interaction ( <i>vs. one interaction</i> ) [80]	o				
	55	High gesture amplitude ( <i>vs. low gesture amplitude</i> ) [78]	–				
Emotional expression	20	Referential eye gaze ( <i>vs. affiliative eye gaze vs. both referential and affiliative eye gaze</i> ) [48]	–				
	33	Happy facial expression ( <i>vs. sad facial expression</i> ) [51]	+				
	33	Happy facial expression ( <i>vs. sad facial expression</i> ) [58]	+				
	8	Positive and negative emotion in language ( <i>vs. neutral language</i> ) [50]			+		
Embodiment	63	Empathic language after neutral language ( <i>vs. neutral language after empathic language</i> ) [83]	+				
	44	Virtual embodiment ( <i>vs. a social robot</i> ) [55]				o	
Environment	40	Virtual embodiment ( <i>vs. voice-only</i> ) [84]	+				
	43	Realistic background ( <i>vs. featureless grey background</i> ) [57]	o				
Combination	8	Personal level physical proximity to screen ( <i>vs. intimate vs. social vs. public level physical proximity to screen</i> ) [85]			o		
	1607	Relational language and nonverbal cues ( <i>vs. no relational language and nonverbal cues</i> ) [62]		+			
	1607	Relational language and nonverbal cues ( <i>vs. no relational language and nonverbal cues</i> ) [72]		+			
	84	Relational language and nonverbal cues ( <i>vs. no relational language and nonverbal cues</i> ) [75]	o			+	+

*N* refers to the number of participants in each study

Referential eye gaze refers to eye gaze focused predominantly on a referent object as opposed to the user



the opposite. This study compared the delivery order of a neutral and empathic virtual therapist in a crossover design and found that rapport significantly decreased with time in both conditions [83]. However, rapport decreased most when experiencing an empathic virtual therapist after a neutral virtual therapist [83]. Reasons for this effect are unclear.

### 3.5.5 Embodiment

Two studies looked at the effect of an ECA's virtual embodiment on ratings of relationship quality [32, 84]. One study found that an ECA received higher rapport ratings in comparison to a voice-only agent in a group decision making context [84]. Another study found no significant difference between an ECA and a physical robot on ratings of working alliance with older adults following 30 days of interaction [32]. This suggests that ECAs may develop better relationships than voice-only agents and similar relationships to social robots, however further research is needed to replicate findings.

### 3.5.6 Environment

Aspects of the virtual environment, including a realistic background and a personal level of physical proximity to the screen, were not found to have any significant effects on rapport or intimacy [57, 85].

### 3.5.7 Combination of Features

A combination of verbal and nonverbal relational cues was shown to improve perceptions of social closeness, working alliance, and relationship quality across three experimental studies [62, 72, 75]. All studies evaluated the same combination of cues, which included empathic language, social dialogue, meta-relational communication, humour, continuity behaviour, including the user's name during a greeting, politeness strategies, and immediacy behaviour such as head nods, eye gaze, closeness to the screen, eyebrow raises, and hand gestures. This combination of relational cues was tested in exercise coaching and museum guiding contexts with ECAs that were humanlike and robotlike in appearance. In all three studies, comparisons were made to ECAs who delivered no relational cues during the interaction. The literature to date suggests that a combination of verbal and nonverbal relational cues shows promise for improving perceptions of relationship quality with ECAs of diverse appearances and application contexts. However, as cues were evaluated together, it is not possible to discern which cues contribute most to improvements in relationship quality.

## 3.5.8 Effect of Design Features on Social Perceptions and Behaviours

This section presents results for question 2 (“What design features are shown to improve social perceptions and behaviours towards ECAs?”). 40 studies evaluated the effect of a design feature on social perceptions and behaviours that could affect relationship quality. 39 social perceptions and behaviours were evaluated across the studies, including trust, engagement, desire to interact again, self-disclosure intimacy and amount, caring, warmth, felt supported, social attraction, and intention to use. 34 unique design features were tested that related to language ( $n=12$ ), behaviour ( $n=8$ ), emotional expression ( $n=5$ ), embodiment ( $n=2$ ), appearance ( $n=4$ ), personality ( $n=1$ ), the virtual environment ( $n=1$ ), and a combination of language and behaviour features ( $n=1$ ). Results are depicted in Table 3.

### 3.5.9 Language

13 studies evaluated the effect of language features on social perceptions and behaviours towards an ECA. Four studies focused on adding social components to language, such as small talk and friendly chat. Small talk was shown to improve perceived trust, knowledge of the user, and success of the interaction for extroverted users, while task-oriented talk received higher ratings from introverted users [86]. Social-oriented plus task-oriented language was shown to increase perceptions of an engaging personality and trustworthiness over task-oriented language alone, however only for older adults with high internet competency [59]. A possible reason for this effect is that older adults with low internet competency reported experiencing more information overload while interacting with an ECA with social-oriented language. In another study, ECAs with friendly language were rated higher in warmth than those with neutral language [66]. Wordiness was found to increase user's self-disclosure as well as improve positive interviewer and interaction perceptions for an ECA deployed in an interviewing context [87].

Several studies investigated the effect of self-disclosure on social perceptions and behaviours. High self-disclosure was found to improve self-disclosure intimacy, ratings of social attraction and presence [73], and the amount of self-disclosure for users willing to disclose a medium amount [74]. However, high self-disclosure was also associated with decreased likeability and self-disclosure for those willing to disclose a low amount [74]. Self-disclosure of humanlike back stories was associated with higher self-disclosure intimacy and amount for users high in anxiety [53]. ECAs with a first-person storytelling perspective were shown to elicit more self-disclosure from users compared to ECAs with a third-person storytelling perspective

**Table 3** Effect of design features on social perceptions and behaviours of ECAs (+ = increase in outcome; # = increase dependent on a user characteristic; – = decrease in outcome; **o** = no effect on the social perception outcome,  $p < .05$ )

Design features	<i>N</i>	Subcomponents	Trust	Engagement	Desire to interact again	Self-disclosure amount	Self-disclosure intimacy	Other social perceptions and behaviours
Language	31	Small talk language ( <i>vs. task-oriented language only</i> ) [86]	+	<b>o</b>	<b>o</b>			Knowledge of user (#); Success of interaction (#); Interest in interaction (o)
	24	Variable language ( <i>vs. no variable language</i> ) [88]		<b>o</b>	<b>o</b>			Caring (o); Helpfulness (o)
	8	Context awareness in language ( <i>vs. no context awareness in language</i> ) [33]	<b>o</b>					
	29	User identification ( <i>vs. no user identification</i> ) [72]		<b>o</b>				
	121	Social-oriented plus task-oriented language ( <i>vs. task-oriented language only</i> ) [59]	+					Engaging personality (#); Perceived information overload (#)
	60	First-person storytelling perspective ( <i>vs. third-person storytelling perspective</i> ) [77]			+			Warmth (#)
	40	Friendly language ( <i>vs. neutral language</i> ) [66]						Social presence (#); Social attraction (#)
	57	High self-disclosure ( <i>vs. low self-disclosure vs. no self-disclosure</i> ) [73]				+		Social attraction (o); Likeability (–)
	171	High self-disclosure ( <i>vs. low self-disclosure vs. no self-disclosure</i> ) [74]				+		
	40	Self-disclosure of humanlike stories ( <i>vs. self-disclosure of computerlike stories</i> ) [53]				+		
	81	Wordiness ( <i>vs. taciturn language</i> ) [87]				+		Positive interviewer perceptions (#); Pleasant interaction (#)
	10	Accurate recall of user information ( <i>vs. no recall vs. inaccurate recall vs. partial recall</i> ) [89]	<b>o</b>					
Behaviour	24	Good timing eye gaze ( <i>vs. bad timing eye gaze vs. static eye gaze</i> ) [90]			+			Thoughtfulness (o)
	72	High eye gaze ( <i>vs. low eye gaze</i> ) [91]				+		Negative partner perceptions (–)
	20	Affiliative eye gaze ( <i>vs. referential eye gaze</i> ) [48]	+	+				

Table 3 (continued)

Design features	N	Subcomponents	Trust	Engagement	Desire to interact again	Self-disclosure amount	Self-disclosure intimacy	Other social perceptions and behaviours
	80	Gesture behaviour present (vs. <i>no gesture behaviour</i> ) [33]						Warmth (o)
	171	High behavioural realism (e.g. breathing, blinking, head nods, posture shifts) (vs. <i>low behavioural realism</i> ) [74]			+			Social attraction (o)
	83	High behavioural realism (e.g. breathing, blinking, head nods, posture shifts) (vs. <i>low behavioural realism</i> ) [82]			+			Person perception (o); Empathy (o); Mutual understanding (o); Mutual awareness (o)
	19	Repeated interaction (vs. <i>one interaction</i> ) [57]						Compassion (#)
	79	Socially responsive behavioural feedback (e.g. head nods, smiles) (vs. <i>no socially responsive behavioural feedback</i> ) [81]						Social connectedness (o)
	40	Mimicry of user facial expression (e.g. nod, brow raise) (vs. <i>no mimicry behaviour</i> ) [71]		o			o	
	40	Expressive facial gestures (e.g. smiling, winking, rolling eyes) (vs. <i>no expressive facial gestures</i> ) [71]		+				
	67	Moderate blinking rate (18 blinks per minute) (vs. <i>low blink rate vs. high blink rate</i> ) [92]						Friendliness (for female ECAs only) (#)
	69	Moderate blinking rate (18 blinks per minute) (vs. <i>low blink rate vs. high blink rate</i> ) [92]						Friendliness (for female ECAs only) (#)
Emotional expression	96	Empathic language and facial expression (vs. <i>no emotion vs. self-directed emotion vs. both empathic and self-directed emotion</i> ) [67]			+			Caring (#); Felt supported (#)

Table 3 (continued)

Design features	N	Subcomponents	Trust	Engagement	Desire to interact again	Self-disclosure amount	Self-disclosure intimacy	Other social perceptions and behaviours
	81	Empathic language and facial expression (vs. <i>neutral language and facial expression</i> ) [93]	+					Intention to use (#); Sociability (#); Enjoyment (#); Usefulness (#); Safety (#); Interaction anxiety (o); Social influence (o)
	33	Happy facial expression (vs. <i>sad facial expression</i> ) [51]			+			
	33	Happy facial expression (vs. <i>sad facial expression</i> ) [58]						Interaction quality (#)
	50	Emotional expression in face (happy, warm, concerned, neutral) and speech (rate, pitch) (vs. <i>neutral face and speech</i> ) [68]	o					Caring (#); Felt supported (o)
	242	Amused and polite smiles (vs. <i>no smile</i> ) [94]						Warmth (#)
Embodiment	48	A face (vs. <i>voice-only vs. text-only</i> ) [65]	+					
	50	A face (vs. <i>voice and text-only</i> ) [59]	+					Perceived social support (#)
	81	A face (vs. <i>text-only</i> ) [93]	+					Intention to use (#); Sociability (#); Enjoyment (#); Usefulness (#); Safety (#); Interaction anxiety (o); Social influence (o)
	72	A face (vs. <i>voice and text-only vs. voice-only vs. text-only</i> ) [95]						Social presence (o)
	44	Virtual embodiment (vs. <i>a social robot</i> ) [32]	o					Social agent rating (o)
	40	Virtual embodiment (vs. <i>voice-only</i> ) [84]	+					
Appearance	72	Female gender (vs. <i>male gender</i> ) [91]						Positive partner perception (#)
	80	Humanlike appearance (vs. <i>robotlike appearance</i> ) [54]						Warmth (o)
	168	Humanlike voice (vs. <i>text-to-speech voice vs. text</i> ) [55]	+					Intention to use (#); Social presence (#)
	72	Text-to-speech voice (vs. <i>text</i> ) [95]						Flow (#)

**Table 3** (continued)

Design features	N	Subcomponents	Trust	Engagement	Desire to interact again	Self-disclosure amount	Self-disclosure intimacy	Other social perceptions and behaviours
Personality	80	Introversion in speech and behaviour (slower speech rate & calm tone, neutral face) (vs. <i>extroversion in speech and behaviour</i> ) [2]		+				
Environment	43	Realistic background (vs. <i>featureless grey background</i> ) [57]						Social attraction (#)
Combination	1607	Relational language and non-verbal cues (vs. <i>no relational language and nonverbal cues</i> ) [62]	+		+			
	1607	Relational language and non-verbal cues (vs. <i>no relational language and nonverbal cues</i> ) [72]	+		+			
	84	Relational language and non-verbal cues (vs. <i>no relational language and nonverbal cues</i> ) [75]			+			

N refers to the number of participants in each study

High eye gaze refers to maintaining eye gaze with the user for most of the conversation

[77]. No effects were observed for variable language [88], context awareness in language [33], calling the user by their name [72], or accurate recall of user information [89] on social perceptions. Evidently, a range of language features have been shown to improve social perceptions and behaviours towards ECAs, however the effects of certain features may be impacted by user characteristics.

### 3.5.10 Behaviour

12 studies evaluated the effect of behavioural features, such as eye gaze, gestures, and nonverbal feedback, on a range of social perceptions and behaviours. Three studies looked at the effect of eye gaze. An ECA with referential eye gaze (head turned towards a map for most of the conversation, briefly glanced at the participant) was rated as less trustworthy and engaging than an ECA with affiliative eye gaze (head turned towards the participant for most of the conversation, briefly glanced at a referential object) and both types of eye gaze [48]. In another study, an ECA with ‘good timing’ eye gaze (informed by observation of human dyads) was shown to elicit a greater amount of self-disclosure in comparison to an ECA with a static gaze and an ECA with ‘bad timing’ eye gaze (the opposite of ‘good timing’ eye gaze patterns) [90]. High eye gaze (which involved maintaining eye contact for most of the conversation) in comparison to low eye gaze (eye contact only a few times while listening) was shown to increase self-disclosure intimacy [91]. However, high eye gaze also increased negative partner perceptions. Findings suggest direct eye gaze is better than little to no direct eye gaze, however too much direct eye gaze may negatively affect user perceptions and behaviours.

Other studies looked at the effect of an ECA’s gestures on social perceptions and behaviour. Two studies focused on behavioural realism, where ECAs showed behaviours like breathing, blinking, posture shifts, and back-channeling and understanding head nods. In both studies, an ECA with high behavioural realism was shown to elicit more self-disclosure from users in comparison to an ECA with low behavioural realism [74, 82]. However, no significant differences were observed in person perception, empathy, mutual understanding, mutual awareness, or social attraction.

Other behavioural features shown to improve social perceptions and behaviour were repeated interactions (increased compassion) [57], expressive facial gestures such as smiling, winking, and rolling eyes (increased engagement) [71], and a moderate blinking rate of 18 blinks per minute for female ECAs (increased friendliness) [84]. No significant effects were found for socially responsive behavioural feedback [57], gesture behaviour [54], or mimicry of user facial expressions [71].

### 3.5.11 Emotion

Six experimental studies investigated the effect of emotional expression on several social perceptions and behaviours including trust, self-disclosure, caring, warmth, felt supported, and interaction quality. Empathic language and facial expressions were shown to improve trust, ratings of agent caring and feelings of support [67], as well as intention to use, sociability, enjoyment, usefulness, and safety in comparison to neutral ECAs [93]. Happy facial expressions resulted in higher amounts of self-disclosure [51], as well as higher ratings of the quality of the interaction, especially for users high in extroversion and neuroticism [58]. Conversely, in the same study, sad facial expressions received higher ratings of interaction quality for users high in conscientiousness [58]. One study found that a combination of polite smiles (smiling while greeting a user) and amused smiles (smiling while telling a riddle) received higher ratings of agent warmth in comparison to polite smiles alone and no smiles [94]. In another study, a virtual nutrition coach with an emotional facial expression (happy, warm, concerned, and neutral) and an emotional voice (speech rate and pitch) was rated as significantly more caring than a virtual coach with neutral expressions [68]. In the same study, no significant differences were observed for agent trust and feelings of support. Overall, the literature suggests that the expression of positive emotion and concern improves social perceptions and behaviours towards ECAs, however responses can differ based on user personality.

### 3.5.12 Embodiment

Several studies evaluated the effect of a face and virtual embodiment on social perceptions and behaviours with mixed results. One study found that an ECA received significantly higher ratings of trust than speech-only and text-only agents on a shopping website [65]. Another study found that for older adults, an ECA with speech and text was perceived as more trustworthy and provided better social support over a speech and text-only agent [59]. An empathic virtual counsellor with a face was rated as significantly more trustworthy, sociable, enjoyable, useful, safe, with a higher intention to use in future in comparison to a text-only counsellor [93]. Similarly, another study found an ECA to be more trustworthy than a voice-only agent in a group decision making context [84]. Evidently, a virtual face or embodiment may help to improve trust and other social perceptions towards a conversational agent. Although, one study found no effect of a virtual face on social presence in comparison to speech-only and text-only agents [95]. Lastly, in another study, no significant difference was found between an ECA and a social robot in engagement and social agent ratings [32], which suggests ECAs may be just as engaging as social



robots. Overall, the literature suggests that ECAs are viewed more positively than text-only agents and may be similarly perceived to social robots.

### 3.5.13 Appearance

Four studies researched the effect of ECA appearance on trust, intention to use, and partner perceptions such as warmth, positivity, and social presence. An ECA with a humanlike appearance did not significantly differ from an ECA with a robotic appearance on perceptions of warmth during an educational task [54]. This suggests that a humanlike appearance may not be necessary to perceive an ECA as warm in personality, however further research is needed. A humanlike voice was shown to improve trust, intention to use, and social presence over a text-to-speech voice and text communication [55]. Another study found that an artificial text-to-speech voice was perceived to have better flow than text-only communication [95]. Results suggest that a humanlike voice may be the most preferable option, however in the absence of a humanlike voice, a text-to-speech voice is preferable to text only. One study looked at the effect of gender on partner perceptions and found that ECAs with female gender were rated higher in positive partner perceptions than male ECAs in a companionship context [91]. Further research is needed to understand the reasons for this effect.

### 3.5.14 Personality

One study investigated the effect of an ECA's personality on trust. In this study, an extroverted ECA (faster speech rate, larger pitch range, frequent smiles, and expansive head gesture) was rated as less trustworthy than an introverted ECA (slower speech rate, calm vocal tone, neutral facial expression, and low head animation) for extroverted participants [2]. There were no significant differences in willingness to trust an extroverted or introverted ECA for introverted participants. This suggests a difference in ECA personality preference by user personality, and further research is needed to understand how other personality traits might affect ECA preferences.

### 3.5.15 Environment

Only one study evaluated the effect of the virtual environment on social perceptions. In this experiment, a realistic background, which was an animated video of an outdoor scene, improved perceptions of social attraction for male users [57]. Whereas for female users, a featureless grey background resulted in higher social attraction towards the ECA. The results of this study suggest a gender difference in virtual environment preferences, however further research

is needed to replicate and understand the reasons for this gender effect.

### 3.5.16 Combination of Features

Relational verbal and nonverbal cues were shown to improve engagement and desire to interact in future across several studies previously described in the relationship quality results section of this review [62, 72, 75].

## 3.6 Risk of Bias Assessment

A risk of bias assessment revealed several key areas in which the quality of included studies could be improved. Included studies could have improved on reporting how randomization procedures were conducted and whether personnel were blinded to participant allocation, as well as provide data or citations on the reliability and validity of subjective self-report measures. The quality of evaluations could have been improved with the use of validated scales over ad hoc single item measures for subjective outcomes in several cases. Aspects that were typically conducted well with a low degree of bias were participant randomization and outcome reporting. However, some studies did not report full statistical information such as effect sizes or mean scores. Overall, the risk of bias for included studies was generally low with some concerns pertaining to measurement quality, and reporting of allocation processes, blinding, and statistics.

## 4 Discussion

### 4.1 Summary of Evidence

This systematic review identified 47 studies from 43 unique publications which evaluated the effect of a design feature on relationship quality and/or social perceptions and behaviours towards an ECA. A range of design features were shown to improve relationships or social perceptions and behaviours towards ECAs. These included virtual embodiment, a humanlike voice, language content (e.g. small talk, self-disclosure, humour), relationship building behaviour (e.g. affiliative eye gaze, cooperation), emotional expression in the face and speech, combinations of relational language and behaviour, and the virtual environment. The effects of design features sometimes differed depending on user characteristics including: gender, personality, technical competency, and level of anxiety.

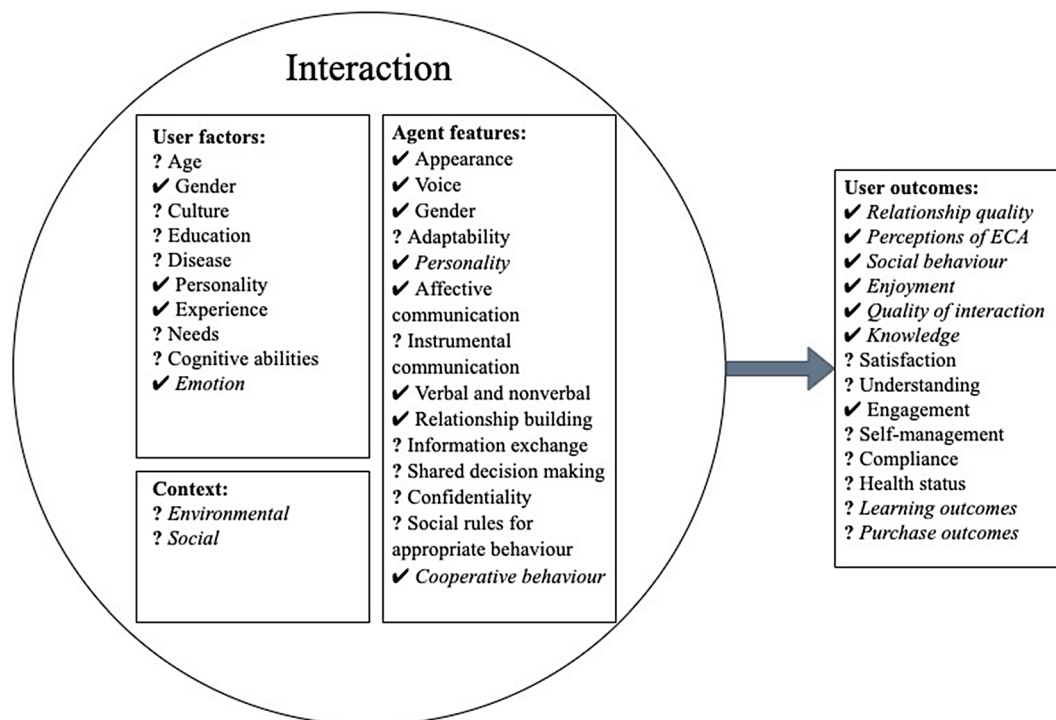
Relationships with ECAs were evaluated across a range of settings including: companionship, healthcare, commercial, and education. While some findings could be applicable across a range of settings, it is important to consider that design features may be perceived differently by users

depending on the setting or particular task for which the ECA is deployed. For example, an ECA that uses humour may be well-received if it is tasked with being a companion, yet badly received if it is discussing sensitive topics such as suicide, sexually transmitted diseases, or debt. There is a need for further research to evaluate which design features are important for improving relationships with ECAs across different use cases.

A wide range of ECAs were evaluated that varied in appearance characteristics and presentation features, however it is important to note the majority of ECAs were humanlike, female adults with white skin colour. Another issue is that although a wide range of design features have been evaluated across the literature, often there are only one or two studies evaluating the effect of a particular feature. More studies are needed to see whether results can be replicated, including with more diverse ECA designs and user populations.

Figure 5 shows a framework for the factors that affect relationship quality, social perceptions and behaviours towards ECAs from the evidence reviewed here. The framework is built on the model of patient-robot communication

[47], but when incorporating the results of this review, it has been transformed into a different figuration and generalised to other settings. Support has been found for the effects of ECA appearance, gender, voice, and personality on user outcomes, but there is a lack of research on adaptability (the ability of the ECA to change behaviour based on user characteristics). Evidence has also been found for effects of user gender, personality, experience (technical competency), and level of anxiety on relationship quality and/or social perceptions. There is a lot of evidence for the effects of verbal behaviours, as well as non-verbal behaviours (facial expressions, eye gaze) on relationship quality, social perceptions and behaviours. There is also evidence for affective communication and relationship building. However, there is a lack of research to date on the effects of information exchange, shared decision making, confidentiality, and appropriate medical behaviour on relationship quality, social perceptions and behaviours towards ECAs. Regarding outcomes, this review focused on relationship quality, social perceptions and behaviours. Other outcomes, such as adherence or health status, were not examined, and could be the focus of future reviews.



**Fig. 5** Summary of evidence to date on ECA features shown to affect user outcomes (framework adapted from the model of robot-patient communication [47]). This framework proposes that factors related to the user, the context, and the agent's features may have interaction or main effects on user outcomes. The evidence for the effects of ECA

features on user outcomes is summarised as follows: ✓ = supportive evidence was found; ? = no evidence to date. (*Italicised* features are new factors compared to the original model of robot-patient communication)

## 4.2 Limitations

This review had several limitations pertaining to the quality of included studies. The risk of bias in included studies was generally low, however some concerns related to the quality of outcome measurement and reporting of statistical information, allocation and blinding procedures arose during a risk of bias assessment. 53% of included articles were published at engineering or computer science conferences where studies were unlikely to be required to meet the same psychometric standards as publications in health or psychology journals (however, engineering conference papers often meet the same peer reviewing standards as engineering or computer science journals). These limitations make it difficult to ascertain how robust findings are.

There were also several possible limitations of this review. While a concerted effort was made to attain literature across computer and health science disciplines, it is possible that some studies may have been missed due to search strategy, database choice, and restricting results to English language only. We also limited the review to ECAs and did not include studies on physical robots. This was because feasible design features are likely to differ based on whether embodiment is virtual or physical (e.g. touch). However, some findings included in this review are likely to generalize to other embodied technologies such as social robots or digital humans. Digital humans are a more recent form of an ECA that incorporate artificial intelligence [96].

## 4.3 Gaps in the Research

This review identified several gaps in the research literature. First, further research is needed on developing quality relationships with ECAs across a broader range of task domains. In healthcare, for example, only nine studies have looked at improving relationships with ECAs. Studies were conducted in the context of mental health counselling, and diet and exercise coaching. More research is needed to understand what design features are important to building quality relationships with ECAs with a broader range of patient populations and applied to more healthcare support tasks (e.g. reminders for medication adherence, delivery of informational support, provision of mindfulness meditation exercises). The unique demands of different patient populations or requirements of various healthcare support tasks could mean that different ECA designs are more or less suitable. More studies looking at specific task domains in education, commerce, and other industries are also needed.

Second, no studies have investigated the effect of an ECA's appearance or personality on the quality of its relationships with users (only social perceptions and behaviours). Moreover, only a small number of studies have looked at the effects of virtual embodiment, emotional

expression, the environment, and a combination of factors on relationships with ECAs. More studies are needed to map new design features that benefit relationships with ECAs and to replicate existing findings.

Third, there is a need for improvement in the quality of research methods used when evaluating the effect of a design feature on relationships with users. Many publications did not use validated scales or did not provide data on scale reliability when measuring subjective self-report outcomes. It is important that future research uses valid and reliable scales to ensure subjective outcomes, like relationship quality, are appropriately measured. There was also a lack of consistency across the research in terms of both outcome selection and measurement of outcomes. In order to adequately compare the effect of different design features, it is important that the research field strives for consistency in what relationship quality outcome is used as well as how this is measured. Future studies could measure rapport, given this was shown to be the most common outcome in the literature and it applies to both professional and personal relationships.

Fourth, this review found a lack of diversity in ECAs used in the literature in terms of age, gender, and skin colour. Of the 72 ECAs that appeared in the literature, 53 were adult age, 38 were female, and 48 were white-skinned. Although there was not a large difference in the ratio of female to male ECAs (38:33), future research could look at using androgynous ECAs. Studies could also study the effect of using ethnically-diverse ECAs of varying ages. Users may feel a greater sense of closeness to ECAs that are demographically similar to them, given closeness in human relationships is affected by the homophily principle [97].

## 5 Conclusion

Overall, this systematic review found that a range of design features can be used to improve relationships with ECAs. Results show that features pertaining to language content, relationship building behaviour, emotional expression, and physical characteristics such as voice type and virtual embodiment may improve relationships with ECAs. Results suggest that there may be differences in the ways people respond to design features based on characteristics such as their gender, personality, technology competency, and social anxiety. There is a need for further research on design features for improving relationships with ECAs across a broader range of use cases and contexts, using robust research methods, with more diverse user populations and ECA designs. ECAs show considerable promise for increasing supply of supportive services, healthcare interventions, and providing companionship to people, but there is a way to go in understanding how to make ECAs effective and engaging for all consumers.

## Compliance with Ethical Standards

**Conflicts of interest** MS is the CEO of Soul Machines (an artificial intelligence company), which supports KL with a PhD stipend, and contracts EB for consultancy work.

## References

- Cassell J, Sullivan J, Churchill E, Prevost S (eds) (2000) Embodied conversational agents. MIT Press, New York
- Liew TW, Tan SM (2016) The effects of positive and negative mood on cognition and motivation in multimedia learning environment. *J Educ Technol Soc* 19(2):104–115
- Tamayo S, Perez-Marin D (2012) An agent proposal for reading understanding: applied to the resolution of maths problems. In: 2012 international symposium on computers in education (SIIE), pp 1–4
- Ryokai K, Vaucelle C, Cassell J (2003) Virtual peers as partners in storytelling and literacy learning. *J Comput Assist Learn* 19(2):195–208
- Lopez V, Eisman EM, Castro JL (2008). A tool for training primary health care medical students: the virtual simulated patient. In 20th IEEE international conference on tools with artificial intelligence, pp 194–201
- Hayashi Y (2015) Influence of social communication skills on collaborative learning with a pedagogical agent: investigation based on the Autism-spectrum quotient. In Proceedings of the 3rd international conference on human-agent interaction, pp 135–138
- Roorda DL, Koomen HMY, Split JL, Oort FJ (2011) The influence of affective teacher-student relationships on students' school engagement and achievement: a meta-analytic approach. *Rev Educ Res* 81(4):493–529
- Cornelius-White J (2007) Learner-centered teacher-student relationships are effective: a meta-analysis. *Rev Educ Res* 77(1):113–143
- McBreen H, Shade P, Jack M, Wyard P (2000) Experimental assessment of the effectiveness of synthetic personae for multimodal e-retail applications. *Proc Fourth Int Conf Autonom Agents* 3(7):39–45
- Matthews A, Anderson N, Anderson J, Jack M (2008) Individualised product portrayals in the usability of a 3D embodied conversational agent in an ebanking scenario. In: Proceedings of the international workshop on intelligent virtual agents, pp 516–517
- Cassell J, Bickmore T, Billinghurst M, Campbell K, Chang K, Vilhjalmsón H, Yan H (1999) Embodiment in conversational interfaces: Rea. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 520–527
- Arafa Y, Mamdani A (2000) Face-to-face interaction with an electronic personal sales assistant. *Proc IEEE Conf Syst Man Cybern* 2:792–797
- Beldad A, Hegner S, Hoppen J (2016) The effect of virtual sales agent (VSA) gender—product gender congruence on product advice credibility, trust in VSA and online vendor, and purchase intention. *Comput Hum Behav* 60:62–72
- Vasiljevs A, Skadina I, Deksnė D, Martins KT, Vira LI (2017) Application of virtual agents for delivery of information services. In: Proceedings of the international scientific conference on new challenges of economic and business development, pp 667–678
- Cyr D, Hassanein K, Head M, Ivanov A (2007) The role of social presence in establishing loyalty in e-service environments. *Interact Comput* 19(1):43–56
- Gardiner PM, McCue KD, Negash LM, Cheng T, White LF, Yinusa-Nyahkoon L, Bickmore TW (2017) Engaging women with an embodied conversational agent to deliver mindfulness and lifestyle recommendations: a feasibility randomized control trial. *Patient Educ Couns* 100(9):1720–1729
- Provoost S, Lau HM, Ruwaard J, Riper H (2017) Embodied conversational agents in clinical psychology: a scoping review. *J Med Internet Res* 19(5):e151
- Bickmore T, Pfeifer L (2008) Relational agents for antipsychotic medication adherence. In: CHI'08 workshop on technology in mental health
- Edwards RA, Bickmore T, Jenkins L, Foley M, Manjourides J (2013) Use of an interactive computer agent to support breastfeeding. *Matern Child Health J* 17(10):1961–1968
- Bickmore TW, Schulman D, Sidner C (2013) Automated interventions for multiple health behaviors using conversational agents. *Patient Educ Couns* 92(2):142–148
- Bickmore TW, Silliman RA, Nelson K, Cheng DM, Winter M, Henault L, Paasche-Orlow MK (2013) A randomized controlled trial of an automated exercise coach for older adults. *J Am Geriatr Soc* 61(10):1676–1683
- Watson A, Bickmore T, Cange A, Kulshreshtha A, Kvedar J (2012) An internet-based virtual coach to promote physical activity adherence in overweight adults: randomized controlled trial. *J Med Internet Res* 14(1):e1
- House J, Kahn R (1985) Measures and concepts of social support. In: *Social support and health*. New York: Academic Press
- Hogan BE, Linden W, Najarian B (2002) Social support interventions: do they work? *Clin Psychol Rev* 22(3):381–440
- Kiecolt-Glaser JK, Gouin JP, Hantsoo L (2010) Close relationships, inflammation, and health. *Neurosci Biobehav Rev* 35(1):33–38
- Uchino BN, Cacioppo JT, Kiecolt-Glaser JK (1996) The relationship between social support and physiological processes: a review with emphasis on underlying mechanisms and implications for health. *Psychol Bull* 119(3):488
- Robinson H, Ravikulan A, Nater UM, Skoluda N, Jarrett P, Broadbent E (2017) The role of social closeness during tape stripping to facilitate skin barrier recovery: preliminary findings. *Health Psychol* 36(7):619
- Heinrichs M, Baumgartner T, Kirschbaum C, Ehlert U (2003) Social support and oxytocin interact to suppress cortisol and subjective responses to psychosocial stress. *Biol Psychiatry* 54(12):1389–1398
- Valtorta NK, Kanaan M, Gilbody S, Ronzi S, Hanratty B (2016) Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. *Heart* 102(3):1009–1016
- Hawkey LC, Cacioppo JT (2010) Loneliness matters: a theoretical and empirical review of consequences and mechanisms. *Ann Behav Med* 40(2):218–227
- Ring L, Barry B, Totzke K, Bickmore T (2013) Addressing loneliness and isolation in older adults: proactive affective agents provide better support. In: 2013 Humaine Association conference on affective computing and intelligent interaction, pp 61–66. IEEE
- Sidner CL, Bickmore T, Nooraie B, Rich C, Ring L, Shayganfar M, Vardoulakis L (2018) Creating new technologies for companionable agents to support isolated older adults. *ACM Trans Interact Intell Syst (Tiis)* 8(3):17
- Bickmore TW, Mauer D, Brown T (2009) Context awareness in a handheld exercise agent. *Pervas Mobile Comput* 5(3):226–235
- Jeong S, Breazeal C, Logan D, Weinstock P (2018) Huggable: the impact of embodiment on promoting socio-emotional interactions for young pediatric inpatients. In: Proceedings of the 2018 CHI conference on human factors in computing systems, ACM, pp 1–13
- Kessler RC, McLeod JD (1985) Social support and mental health in community samples. Academic Press, New York



36. Uchino BN (2006) Social support and health: a review of physiological processes potentially underlying links to disease outcomes. *J Behav Med* 29(4):377–387
37. Ring L, Shi L, Totzke K, Bickmore T (2015) Social support agents for older adults: longitudinal affective computing in the home. *J Multimodal User Interfaces* 9(1):79–88
38. Broadbent E, Garrett J, Jepsen N, Ogilvie VL, Ahn HS, Robinson H, Peri K, Kerse N, Rouse P, Pillai A, MacDonald B (2018) Using robots at home to support patients with chronic obstructive pulmonary disease: pilot randomized controlled trial. *J Med Internet Res* 20(2):e45
39. Joranson N, Pedersen I, Rokstad AMM, Ihlebaek C (2015) Effects of symptoms of agitation and depression in persons with dementia participating in robot-assisted activity: a cluster-randomized controlled trial. *J Am Med Dir Assoc* 16(10):867–873
40. Chipidza FE, Wallwork RS, Stern TA (2015) Impact of the doctor-patient relationship. In: *The primary care companion for CNS disorders*, vol 17, no 5
41. Stewart MA (1995) Effective physician-patient communication and health outcomes: a review. *CMAJ Can Med Assoc J* 152(9):1423
42. Kaplan SH, Greenfield S, Ware Jr, JE (1989) Assessing the effects of physician-patient interactions on the outcomes of chronic disease. *Med Care* S110–S127
43. Roberts CS, Cox CE, Reintgen DS, Baile WF, Gibertini M (1994) Influence of physician communication on newly diagnosed breast patients' psychologic adjustment and decision-making. *Cancer* 74(S1):336–341
44. Schillinger D, Bindman A, Wang F, Stewart A, Piette J (2004) Functional health literacy and the quality of physician-patient communication among diabetes patients. *Patient Educ Couns* 52(3):315–323
45. Thompson L, McCabe R (2012) The effect of clinician-patient alliance and communication on treatment adherence in mental health care: a systematic review. *BMC Psychiatry* 12(1):87
46. Alexander JA, Hearld LR, Mittler JN, Harvey J (2012) Patient-physician role relationships and patient activation among individuals with chronic illness. *Health Serv Res* 47(3):1201–1223
47. Broadbent E, Johanson D, Shah J (2018) A new model to enhance robot-patient communication: applying insights from the medical world. In: Ge S et al (eds) *Social robotics. ICSR 2018. Lecture notes in computer science*, vol 11357. Springer, Cham, pp 308–317. [https://doi.org/10.1007/978-3-030-05204-1\\_30](https://doi.org/10.1007/978-3-030-05204-1_30)
48. Pejsa T, Andrist S, Gleicher M, Mutlu B (2015) Gaze and attention management for embodied conversational agents. *ACM Trans Interact Intell Syst (Tiis)* 5(1):3
49. Kulms P, Kopp S, Krämer NC (2014) Let's be serious and have a laugh: can humor support cooperation with a virtual agent?. In: *International conference on intelligent virtual agents*, Springer, Cham, pp 250–259
50. Partala T, Surakka V, Lahti J (2004) Affective effects of agent proximity in conversational systems. In: *Proceedings of the third Nordic conference on Human-computer interaction*, pp 353–356. ACM
51. Cerekovic A, Aran O, Gatica-Perez D (2016) Rapport with virtual agents: what do human social cues and personality explain? *IEEE Trans Affect Comput* 8(3):382–395
52. Romero OJ, Zhao R, Cassell J (2017) Cognitive-inspired conversational-strategy reasoner for socially-aware agents. In: *IJCAI*, pp 3807–3813
53. Kang SH, Gratch J, Wiederhold BK, Riva G (2012) Socially anxious people reveal more personal information with virtual counselors that talk about themselves using intimate human back stories. *Ann Rev Cyber Telemed* 181:202–207
54. Bergmann K, Eyssel F, Kopp S (2012) A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time. In: *International conference on intelligent virtual agents*, pp 126–138. Springer, Berlin, Heidelberg
55. Qiu L, Benbasat I (2009) Evaluating anthropomorphic product recommendation agents: a social relationship perspective to designing information systems. *J Manag Inf Syst* 25(4):145–182
56. Shamekhi A, Liao QV, Wang D, Bellamy RK, Erickson T (2018) Face Value? Exploring the effects of embodiment for a group facilitation agent. In: *Proceedings of the 2018 CHI conference on human factors in computing systems*, p 391. ACM
57. Kang SH, Phan T, Bolas M, Krum DM (2016) User perceptions of a virtual human over mobile video chat interactions. In: *International conference on human-computer interaction*, Springer, Cham, pp 107–118
58. Cerekovic A, Aran O, Gatica-Perez D (2014) How do you like your virtual agent? Human-agent interaction experience through nonverbal features and personality traits. In: *International workshop on human behavior understanding*. Springer, Cham, pp 1–15
59. Chattaraman V, Kwon WS, Gilbert JE, Ross K (2019) Should AI-Based, conversational digital assistants employ social-or task-oriented interaction style? A task-competency and reciprocity perspective for older adults. *Comput Hum Behav* 90:315–330
60. Morry MN, Reich T, Kito M (2010) How do I see you relative to myself? Relationship quality as a predictor of self- and partner-enhancement within cross-sex friendships, dating relationships, and marriages. *J Soc Psychol* 150(4):369–392
61. Dush CMK, Amato PR (2005) Consequences of relationship status and quality for subjective well-being. *J Personal Soc Relationsh* 22(5):607–627
62. Bickmore T, Pfeifer L, Schulman D (2011) Relational agents improve engagement and learning in science museum visitors. In: *International workshop on intelligent virtual agents*, Springer, Berlin, Heidelberg, pp 55–67
63. Tickle-Degnen L, Rosenthal R (1990) The nature of rapport and its nonverbal correlates. *Psychol Inq* 1(4):285–293
64. Allison T, Puce A, McCarthy G (2000) Social perception from visual cues: role of the STS region. *Trends Cognit Sci* 4(7):267–278
65. Alotaibi MB, Rigas DI (2012) Fostering the user interface acceptance in customer relationship management: a multimedia-aided approach. In: *2012 ninth international conference on information technology-new generations*, pp 796–801. IEEE
66. Griffiths S, Eyssel FA, Philippson A, Pietsch C, Wachsmuth S (2015) Perception of artificial agents and utterance friendliness in dialogue. In: *Proceedings of the 4th international symposium on new frontiers in human-robot interaction at the AISB convention 2015*
67. Brave S, Nass C, Hutchinson K (2005) Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *Int J Hum Comput Stud* 62(2):161–178
68. Creed C, Beale R (2012) User interactions with an affective nutritional coach. *Interact Comput* 24(5):339–350
69. Dobrian F, Awan A, Joseph D, Ganjam A, Zhan J, Sekar V, Stoica I, Zhang H (2013) Understanding the impact of video quality on user engagement. *Commun ACM* 56(3):91
70. Doherty K, Doherty G (2018) Engagement in HCI: conception, theory and measurement. *ACM Comput Surv* 51(5):1–39
71. Stevens CJ, Pinchbeck B, Lewis T, Luerksen M, Pfitzner D, Powers DM, Gibert G (2016) Mimicry and expressiveness of an ECA in human-agent interaction: familiarity breeds content! *Comput Cognit Sci* 2(1):1
72. Bickmore TW, Vardoulakis LMP, Schulman D (2013) Tinker: a relational agent museum guide. *Auton Agent Multi-Agent Syst* 27(2):254–276
73. Kang SH, Gratch J (2011) People like virtual counselors that highly-disclose about themselves. *Ann Rev Cyber Telemed* 167:143–148

74. Kang SH, Gratch J (2014) Exploring users' social responses to computer counseling interviewers' behavior. *Comput Hum Behav* 34:120–130
75. Bickmore TW, Picard RW (2005) Establishing and maintaining long-term human-computer relationships. *ACM Trans Comput Hum Interact* 12(2):293–327
76. Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia. [www.covidence.org](http://www.covidence.org)
77. Gilani SN, Sheetz K, Lucas G, Traum D (2016) What kind of stories should a virtual human swap? In: International conference on intelligent virtual agents. Springer, Cham, pp 128–140
78. Novick D, Gris I, Camacho A, Rayon A, Gonzalez T (2017) Bigger (Gesture) isn't always better. In: International conference on human-computer interaction. Springer, Cham, pp 609–619
79. Hoegen R, Van Der Schalk J, Lucas G, Gratch J (2018) The impact of agent facial mimicry on social behavior in a prisoner's dilemma. In: Proceedings of the 18th international conference on intelligent virtual agents. ACM, pp 275–280
80. Novick D, Gris I (2014) Building rapport between human and ECA: a pilot study. In: International conference on human-computer interaction. Springer, Cham, pp 472–480
81. Krämer NC, Lucas G, Schmitt L, Gratch J (2018) Social snacking with a virtual agent—On the interrelation of need to belong and effects of social responsiveness when interacting with artificial entities. *Int J Hum Comput Stud* 109:112–121
82. Von der Pütten AM, Krämer NC, Gratch J, Kang SH (2010) It doesn't matter what you are! Explaining social effects of agents and avatars. *Comput Hum Behav* 26(6):1641–1650
83. Ranjartabar H, Richards D, Bilgin A, Kutay C (2019) First impressions count! The role of the human's emotional state on rapport established with an empathic versus neutral virtual therapist. In: IEEE transactions on affective computing
84. Shamekhi A, Liao QV, Wang D, Bellamy RK, Erickson T (2018) Face Value? exploring the effects of embodiment for a group facilitation agent. In: Proceedings of the 2018 CHI conference on human factors in computing systems. ACM, p 391
85. Partala T, Surakka V, Lahti J (2004) Affective effects of agent proximity in conversational systems. In: Proceedings of the third Nordic conference on Human-computer interaction. ACM, pp 353–356
86. Bickmore T, Cassell J (2001) Relational agents: a model and implementation of building user trust. In: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, pp 396–403
87. Von der Pütten A, Hoffmann L, Klatt J, Krämer NC (2011) Quid pro quo? Reciprocal self-disclosure and communicative accommodation towards a virtual interviewer. In: International workshop on intelligent virtual agents. Springer, Berlin, Heidelberg, pp 183–194
88. Bickmore T, Schulman D (2009) A virtual laboratory for studying long-term relationships between humans and virtual agents. In: Proceedings of The 8th international conference on autonomous agents and multiagent systems, Vol 1, pp 297–304. International Foundation for Autonomous Agents and Multiagent Systems
89. Richards D, Bransky K (2014) ForgetMeNot: what and how users expect intelligent virtual agents to recall and forget personal conversational content. *Int J Hum Comput Stud* 72(5):460–476
90. Andrist S, Mutlu B, Gleicher M (2013) Conversational gaze aversion for virtual agents. In: International workshop on intelligent virtual agents. Springer, Berlin, Heidelberg, pp 249–262
91. Kulms P, Krämer NC, Gratch J, Kang SH (2011) It's in their eyes: a study on female and male virtual humans' gaze. In: International workshop on intelligent virtual agents. Springer, Berlin, Heidelberg, pp 80–92
92. Takashima K, Omori Y, Yoshimoto Y, Itoh Y, Kitamura Y, Kishino F (2008) Effects of avatar's blinking animation on person impressions. In: Graphics interface conference, pp 169–176. ACM
93. Lisetti C, Amini R, Yasavur U, Riske N (2013) I can help you change! an empathic virtual agent delivers behavior change health interventions. *ACM Trans Manag Inf Syst* 4(4):19
94. Ochs M, Pelachaud C, Mckeown G (2017) A user perception-based approach to create smiling embodied conversational agents. *ACM Trans Interact Intell Syst* 7(1):4
95. Qiu L, Benbasat I (2005) Online consumer trust and live help interfaces: the effects of text-to-speech voice and three-dimensional avatars. *Int J Hum Comput Interact* 19(1):75–94
96. Sagar M, Bullivant D, Robertson P, Efimov O, Jawed K, Kalarot R, Wu T (2014) A neurobehavioural framework for autonomous animation of virtual human faces. In: Proceedings of the SA'14 SIGGRAPH Asia workshop on autonomous virtual humans and social robot for tele-presence. ACM
97. McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: homophily in social networks. *Ann Rev Sociol* 27(1):415–444

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Kate Loveys** received her B.A. in Psychology and M.Sc. in Health Psychology (Hons) from the University of Auckland, New Zealand. Kate is currently a Ph.D. Candidate in Health Psychology at the Department of Psychological Medicine, the University of Auckland, New Zealand. Her research interests are in technology and psychosomatic medicine, with a particular interest in developing engaging digital humans for delivering psychological therapies that benefit physical and mental health outcomes.

**Gabrielle Sebaratnam** received her B.Sc. in Physiology and GradDipSci in Psychology from the University of Auckland, New Zealand. Gabrielle is currently working towards a Masters in Health Psychology at the Department of Psychological Medicine, the University of Auckland. Her research interests include psychosomatic medicine, with a specific interest in surgical anxiety and how technology can be integrated into healthcare delivery to promote physical and psychological wellbeing.

**Mark Sagar** received his B.Sc. and Ph.D. in Engineering from the University of Auckland, New Zealand, and completed his postdoctoral fellowship at Massachusetts Institute of Technology. Mark is currently an Associate Professor and Director at the Laboratory for Animate Technologies at the Auckland Bioengineering Institute, the University of Auckland, New Zealand, and CEO at Soul Machines Ltd, an artificial intelligence company that specializes in building digital humans. Mark's research interests are in creating cognitive architectures, interactive autonomously animated systems and facial animation. He has received two scientific and engineering Academy awards in 2010 and 2011 for his pioneering work in computer-generated faces.

**Elizabeth Broadbent** received her B.E. (Hons) in Electrical and Electronic Engineering from the University of Canterbury, Christchurch, New Zealand, and her M.Sc. and Ph.D. degrees in Health Psychology from the University of Auckland, New Zealand. Elizabeth is currently a Professor in Health Psychology at the Faculty of Medical and Health Sciences, the University of Auckland. Her research interests include the use of robots and digital humans to improve health outcomes, and the emotional connections that people form with robots/digital humans.