

Department of Education
University of Oxford

M.Sc. Dissertation

Changing Nature of Childhood Environments: Exploring Children's Ontologies of Digital Voice Assistants

Author	Janik Festerling (Student Investigator)
College Affiliation	St. Antony's College
Degree Programme	M.Sc. in Education (Research Design & Methodology)
Dissertation Supervisor	Professor Iram Siraj (Principal Investigator)
Date of Submission	8th August 2019
Word count	19,693 (according to examination regulations)

It’s another part of the Copernican story: We used to think Earth was the center of the universe. Now we think we’re special because we have intelligence and nothing else does.

Stephen Wolfram[†]

[†] Quoted from: Wolfram, S. (2019). Artificial Intelligence and the Future of Civilization. In J. Brockman (Ed.), *Possible Minds: Twenty-Five Ways of Looking at AI* (pp. 266–284). New York, NY, US: Penguin Press.

Abstract

At its very core, this exploratory mixed methods inquiry investigated to what extent the growing presence of digital voice assistants, such as Alexa, Siri or the Google Assistant, might introduce changes to today's childhood environments which translate into consequences that go far beyond their immediate purposes. In particular, based on an integrative review of different academic perspectives, especially related to children's cognitive development and the formation of so-called 'ontological belief systems', the inquiry explored the presence of Alexa et al. within today's childhood environments as well as children's engagement with and ontological perception of such devices. In order to elicit these insights, the inquiry partly relied on self-developed data collection techniques, such as the 'Human-Centred Turing Test Game'. The integrative analysis of children's focus group data (n=26, age range = 6 to 10 years) and parental survey data (n=68), which were collected non-randomly in a German municipal elementary school population, resulted in two main empirical claims as the basis for future research: firstly, children's engagement with digital voice assistants is not always a means to an end but sometimes an end in itself, therefore suggesting that these devices might be able to establish themselves as independent engagement partners during childhood and beyond. Secondly, since children's ontological perceptions of Alexa et al. firmly oscillate between distinct ontological domains, one could suspect that these devices leave unprecedented marks on cognitive development trajectories.

Contents

List of Abbreviations I

List of Figures II

List of Tables III

Overview of Key Terms IV

Introduction 1

1. Literature Review & Research Gap 4

1.1 Empirical Insights into Child-Machine Engagements 4

1.2 Insights into the Nature of Human Anthropomorphism 6

1.3 Empirical & Theoretical Aspects of Human Cognitive Development 7

1.3.1 Possible Minds: The Philosophical Nature of Intelligent Machines 7

1.3.2 The Application of ToM in the Context of Intelligent Machines 9

1.3.3 Ontological Belief Systems: How We Make Sense of the World 10

1.3.4 The Ontological Nature of DVAs: Marginality vs. Category in Its Own Right 11

1.4 Research Gap & Research Questions 14

2. Methodology & Research Design 17

2.1 Methodological Foundations 17

2.2 Research Design Choice 18

2.2.1 Sampling Procedure 20

2.2.2 Research Procedure 22

2.3 Methods of Data Collection 24

2.3.1 Collection of Verbal & Observational Data through Focus Groups 24

2.3.2 Collection of Parental Survey Data through Online Questionnaire 31

2.4 Methods of Data Analysis 35

2.4.1 Qualitative Method: Thematic Analysis 35

2.4.2 Quantitative Method: Descriptive Statistics 36

2.5 Reflexivity & Rigour (1): The Design Quality Dimension 37

2.6 Ethical Considerations 40

3. Results & Discussion 42

3.1 Results of Data Analysis 42
3.1.1 Mixed Methods Survey Results 43
3.1.2 Mixed Methods Focus Group Results 50
3.2 Discussion of Results 61
3.3 Reflexivity & Rigour (2): The Interpretive Dimension 65
Conclusion 66
References VI
Appendix XXIII
Statement in Lieu of an Oath LXX

List of Abbreviations

AI	Artificial Intelligence
CUREC	Central University Research Ethics Committee [at the University of Oxford]
DREC	Departmental Research Ethics Committee [at the Department of Education at the University of Oxford]
DVA	Digital Voice Assistant
HI	Human Intelligence
HCTTG	Human Centred Turing Test Game
MI	Machine Intelligence
NOCH	New Ontological Category Hypothesis
SD	Standard Deviation
SI	Student Investigator
ToM	Theory of Mind
TT	Turing Test

List of Figures

Figure 1. Exemplification of an ontological framework 11

Figure 2. DVAs' embeddedness in the ontological framework according to the NOCH 13

Figure 3. Overview of methods and means of triangulation 19

Figure 4. Chronological overview of research procedure 23

Figure 5. Final set-up of focus group session 1 26

Figure 6. Design details of the TT and the HCTTG 30

Figure 7. Presence of DVAs within childhood environments in the total sample..... 45

Figure 8. Domestic DVA-environments and usage patterns in the group of DVA-users 46

Figure 9. General and DVA-specific engagement with digital technologies and media..... 48

Appendix XIV. Reasons for refraining from using DVAs among non-users..... LXIV

List of Tables

Table 1. Research on children's engagement with various types of intelligent machines 5

Table 2. Children's engagement with DVAs and observed probing schemes 15

Table 3. Overview of research questions 16

Table 4. Detailed procedure of focus group session 1 27

Table 5. Detailed procedure of focus group session 2 29

Table 6. Overview of online questionnaire used in parental survey 34

Table 7. Overview of sample sizes across the nested structure of the data 43

Table 8. Firm beliefs about ontological hierarchy: 'HI ≥ MI' 53

Appendix I. Excursus: Conceptual parallels between the nature of HI and MI XXIII

Appendix II. Detailed description of research procedure XXIV

Appendix III. Checklist to use online questionnaires in a survey design XXVI

Appendix IV. Detailed structure of online questionnaire used in parental survey XXVIII

Appendix V. CUREC documents and approval confirmation XL

Appendix VI. Frequencies of demographical characteristics in parental survey LIV

Appendix VII. Frequencies of other personal characteristics in parental survey LV

Appendix VIII. Summary statistics of attitudinal survey instruments LVI

Appendix IX. Internal reliability of attitudinal survey instruments LVI

Appendix X. Correlation among survey instruments LVII

Appendix XI. Frequencies of DVA-ownership and usage in parental survey LVIII

Appendix XII. Mean comparison of attitudinal survey instruments LIX

Appendix XIII. Code book of thematic analysis applied to parental survey LX

Appendix XV. Code book of thematic analysis applied to focus group data LXV

Overview of Key Terms

Key Term	Definition & Examples	Reference
Artefact	Any humanly made entity (e.g. machines, books, clothes, furniture, art)	(-)
Digital Voice Assistant	Any commercially available voice-user interface that is based on digital technology and primarily designed to serve users' assistive needs (e.g. Amazon's '*Alexa'*, Apple's '*Siri',* Google's '*Google Assistant'*)	(-)
Engagement	Any form of individual participation or involvement in an interactive context, including behavioural, emotional and cognitive aspects	Azevedo (2015), Fredricks, Blumenfeld, & Paris (2004)
Entity	Any perceived subject or object of distinct existence (e.g. humans, animals, other living organisms, natural objects, artefacts, machines)	(-)
Intelligent machine	Any machine that can behave intelligently from the speculator's perspective, regardless of its inherent technological sophistication (e.g. remotely controlled or completely autonomous)	Turing (1950)
Machine	Any technological system that is embedded in a humanly made entity (e.g. computers, smartphones, robots, mechanical apparatuses)	(-)
Machine Intelligence	Capacity of machines to pursue tasks which would require human intelligence if we had to perform them ourselves (synonym: 'Artificial Intelligence')	Kurzweil (1990)
Human Intelligence	Any domain of human nature that draws on the innate or acquired capacities of our brain	(-)
Ontology	*Developmental context*: Cognitive framework that comprises the entirety of an individual's mental concepts and schemes, therefore forming a system of boundaries that allows for a basic categorisation of entities along the lines of their perceived features and attributes	Gaudiello, Lefort, & Zibetti (2015)
	Philosophical context: Scientifically-grounded paradigms that specify theories regarding the nature of being (i.e. what the reality we perceive consists of in its essence)	Simons (2009)

Ontological Category	Foundational distinctions between the broadest classes of entities that humans acquire throughout their development in order to make sense of the world around them	Severson & Carlson (2010)
Principal Investigator	Academic supervisor of the student investigator	(-)
Student Investigator	Author of this text and conductor of the empirical inquiry	(-)
Voice-user interfaces	Engagement points between humans and machines where both entities communicate with each other via spoken language and dialog logic	Cohen, Giangola, & Balogh (2004)

Introduction

From a deterministic point of view, the accelerating course of technological progress in the 21st century can be characterised by Schwab's (2017) widely-known concept of the '*Fourth Industrial Revolution'*: according to this notion, humanity undergoes an inevitable era of technologically enabled transformation that is driven by the large-scale digitisation of society as a whole.[1] Consequently, the formerly *distinct ontological spheres* of the physical, biological world *blur* together through revolutionary digital melting pots in the technological domains of robotics, big data analytics, nanotechnology, sensor technology, bioengineering, or, most importantly, *Machine Intelligence* (MI), which generally refers to the creation of machines that can pursue tasks which would require *Human Intelligence* (HI) if we had to perform them ourselves (Kurzweil, 1990).

At the same time, from a more historical perspective, '*Kranzberg's Law'* states that the nature of technological innovation has always implied "environmental, social, and human consequences that go far beyond the immediate purposes of the technical devices and practices themselves" (Kranzberg, 1986, p.545). For instance, the introduction of the mechanical clock in the 14th century went far beyond the simple purpose of timekeeping by "disassociat[ing] time from human events" and "creat[ing] the belief in an independent world of mathematically measurable sequences" (Mumford, 2010, p.15). Similarly, the invention of the printing press in the 15th century lead to a paradigm shift from the primacy of religion and faith to the primacy of reason and knowledge (Kissinger, 2018). So in line with these illustrations of Kranzberg's Law, it seems reasonable to suspect that similar to the commercial fruits of the 19th century Industrial Revolution, which introduced technological changes to family and childhood environments with unexpected but nevertheless far-reaching consequences (cf. Cowan, 1976), a Fourth Industrial Revolution could have equivalent implications today …

The rise of humanoid voices

But on a daily basis, how does the ongoing large-scale digitisation of society – which, after all, is heavily driven by the corporate tech-quartet of Amazon, Alphabet, Apple and

[1] In contrast to Schwab's (2017) notion of technological determinism – which is also used by other popular authors such as Lee (2018) or Harari (2018) – some academics have provided alternative theoretical accounts of today's societal transformations, such as Webster's (2006) 'Information Society' or Castells' (2004) 'Network Society', which challenge the deterministic narrative that technology *determines* change in all other domains of society – and *not* vice versa. However, due to reasons of written space, an integrative consideration of these alternative viewpoints is beyond the scope of this discussion.

Microsoft – materialise in our cosy family and childhood environments? Well, you could simply ask '*Alexa'*, the '*Google Assistant'*, '*Siri',* or '*Cortana'*, the quartet's humanoid '*Digital Voice Assistants'* (DVAs), which are already present on a massive scale. As we speak, Amazon's Alexa, which was only introduced in 2015, is operating over 100 million times across the globe, while Google has recently announced the deployment of its DVA on over one billion devices (Wiggers, 2019).[2]

From a technological perspective, this rise of DVAs has been enabled by several interdependent technological developments, such as 'machine learning' and 'natural language processing' in the realm of MI (Lecun, Bengio, & Hinton, 2015), or the so-called 'Internet of Things', which refers to the emergence of hardware networks in which all devices are connected to each other via cloud computing (Mazhelis, Luoma, & Warma, 2012). But it would be very misleading to simply consider DVAs as sophisticated products or commodities that serve our needs – and sometimes annoy or even amuse us with their moronic abilities. Instead, from a business perspective, these data-collection hubs are not only designed to become our trusted access portals to information networks, but also omnipresent voice-user interfaces that allow us to engage with entire commercial ecosystems and platforms (Dawar, 2018). Nevertheless, at this point, the MI-driven advancements of DVAs are still seconded by extensive manual human labour (e.g. transcribing and annotating users' voice data that in turn are supposed to further improve the underlying algorithms; Day, Turner, & Drozdiak, 2019). Sceptics of DVAs in particular – or even of MI in general – often ridicule this reliance on manual human labour by stating that MI "is as 'fully-automated' as the Great and Powerful Oz was in that famous scene from the classic film, where Dorothy and friends realize that the great wizard is simply a man manically pulling levers from behind a curtain" (Gray & Suri, 2017, para. 4). Indeed, the role of human manual labour behind the 'MI-curtain' remains an issue with low user awareness. However, metaphorically speaking, one should also acknowledge that even the prompter in the theatre does not necessarily change the audience's perception of the actor's performance, because what counts is the overall magic of the experience. Nevertheless, we shall return to this issue in the subsequent discussion.[3]

[2] Since early 2018, Facebook, the fifth global heavyweight in terms of market capitalisation (PwC, 2018), has also been developing a DVA, but its release date as well as potential applications remain unknown up to this point (Rodriguez, 2019).

[3] Ever since the introduction of Alexa et al., the use of female voices has been heavily criticised on the grounds of sexism (e.g. Hempel, 2015; Nass & Brave, 2005; UNESCO, 2019). However, due to reasons of written space, an integrative consideration of this issue is beyond the scope of this discussion.

But more importantly, from a developmental perspective, the presence of Alexa et al. should also be considered as being systematically different from smart toys (e.g. 'Anki Cozmo') or child-specific DVAs (e.g. 'Mattel Aristotle'), because such product designs are limited to applications during infancy and childhood, and therefore children eventually grow out of their usage at some point. In contrast, Alexa et al. are not limited to specific age ranges, since they are designed as universal personal assistants that could – in theory – accompany you throughout life. To stress this point even further, it would also be misleading to limit the scope of DVAs to private domestic environments: DVAs can already be seamlessly present on stationary as well as mobile devices (e.g. smart speakers, smartphones, car entertainment systems), while business development teams across the globe are also trying to introduce DVAs to the professional world (e.g. Amazon Web Services, 2018). For instance, renowned clinical institutions, such as the Boston Children's Hospital, already explore how commercial DVAs could be used for patient services, clinical processes and even medical diagnostics (Ross, 2019; Small, Nigrin, Churchwell, & Brownstein, 2018).

Present study: assessing the historical momentum of DVAs

So, what could one possibly conclude from DVAs, these concrete manifestations of the Fourth Industrial Revolution which populate a growing number of households across the globe? Although we as adults might take a rather sovereign and firm stance towards these recently introduced humanoid voice-user interfaces, it might be different from the perspective of today's children: based on their engagement with DVAs – which can already emulate the once uniquely human domains of cognition and speech, therefore potentially blurring some of the formerly distinct ontological spheres as Schwab (2017) and others predict – this and future generations might develop unprecedented perceptions of such intelligently behaving machines.

So, at its very core, and inspired by the principle of Kranzberg's Law, this inquiry investigates to what extent DVAs might constitute a profound technological application with historical momentum by introducing changes to today's childhood environments which translate into consequences that go far beyond their immediate purposes. This ambitious empirical approach follows three steps: section (**1**) of this paper briefly reviews the empirical as well as theoretical underpinnings of the given inquiry. Based on the research questions derived under (**1**), section (**2**) discusses all methodological considerations related to the empirical investigation, before results are presented and discussed in section (**3**).

1. Literature Review & Research Gap

When it comes to children's engagement with DVAs as well as their perceptions of these intelligently behaving machines, a conglomerate of academic literature must be worked through, beginning with the most applied domain of child-machine engagement (**1.1**) and culminating in the most abstract domain of children's ontological belief systems (**1.3**).[4] The overall aim of this section is to connect these different perspectives in a coherent argument in order to form this inquiry's conceptual starting point (**1.4**).

1.1 Empirical Insights into Child-Machine Engagements

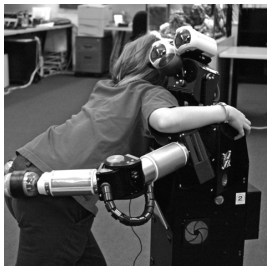
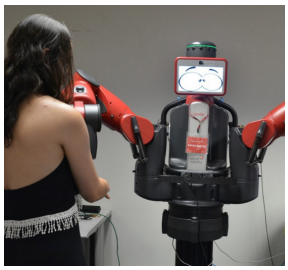
A comprehensive body of empirical literature has investigated how children engage with machines (e.g. humanoid or animalistic robots) that were endowed with varying degrees of technological capabilities in order to make them *appear* intelligent from children's perspective (visual overview in **Table 1**). For instance, in their experimental longitudinal study, Tanaka, Cicourel, & Movellan (2007) showed how infants developed social bonds with a humanoid, non-conversational robot that were similar to the infants' peer-to-peer engagements. Ku, Choi, Lee, Jang, & Do (2018) found additional evidence that infants engaged in less abusive behaviours with an animalistic, non-conversational robot if its physical reactions conveyed the possession of sensitivity and emotions. Furthermore, research that used conversational robots with voice-user interfaces revealed even stronger engagement patterns: for instance, Shiomi, Kanda, Howley, Hayashi, & Hagita (2015) placed a humanoid, conversational robot in an elementary school and showed how children engaged with it in a very social way. Lastly, recent experimental research by Vollmer, Read, Trippas, & Belpaeme (2018) found that robots were even able to induce 'peer-pressure' on children, because their participants between the age of 7 and 9 had a systematic tendency to follow erroneous answers that were given in unison by a group of humanoid robots, while adults were found to be rather resistant. Although the authors could not directly compare participants' conforming behaviour to a scenario with human peers, their findings still support the general notion of strong social engagement patterns that can emerge between children and intelligently behaving machines.

While all the research settings outlined above use robots with humanoid or animalistic physical appearances, the notion of children's engagement with intelligent machines seems

[4] In the context of the given study, the term *engagement* is broadly defined as any form of participation or involvement in an interactive context, including *behavioural*, *emotional* and *cognitive* aspects (cf. Azevedo, 2015; Fredricks, Blumenfeld, & Paris, 2004).

to extend naturally to DVAs, which are "inherently a *socially interactive* device" (p.2854, emphasis in orginal), because users must engage with the machine in a social manner in order to make use of its assistive capabilities (Purington, Taft, Sannon, Bazarova, & Taylor, 2017). In fact, exploratory research based on natural observations in household environments supports this idea of DVAs' social nature, because they become well-integrated devices of daily family life, for instance by providing interactive games to the whole family or by assisting with simple tasks (Porcheron, Fischer, Reeves, & Sharples, 2018). When it comes to child-DVA engagements in particular, children probe and explore various different dimensions of DVAs' perceived intelligence or pre-programmed identities (Druga, 2018; Lovato & Piper, 2015), which are discussed in further detail below (**1.4**).

Table 1. Research on children's engagement with various types of intelligent machines

			
Kahn, Kanda, et al. (2012)	Yarosh et al. (2018)	Ku et al. (2018)	Castro-González et al. (2018)
			
Wood et al. (2017)	Kahn et al. (2006)	Shiomi et al. (2015)	Tanaka et al. (2007)

Notes. Table shows how children across various research contexts and age groups were exposed to different types of intelligently behaving machines (i.e. humanoid robots, animalistic robots, dolls) and how they engaged with them.

Sources. As indicated by references in the table.

In sum, empirical evidence in the area of child-machine engagements suggests that the nature of children's engagements with intelligent machines, such as humanoid and animalistic robots or DVAs, can have various characteristics, such as haptic (e.g. perception of machines' physical reactivity and sensitivity) or conversational experiences (e.g. probing machines' intelligence or identities), and can yield to patterns similar to the social nature of children's intra-human engagements. But what is known about the potential underlying mechanisms of this phenomenon?

1.2 Insights into the Nature of Human Anthropomorphism

As Reeves & Nass (1998) noted in their widely-cited work, "individuals' interactions with computers, […] are fundamentally social and natural, just like interactions in real life" (p.5). This concept of so-called '*anthropomorphism',* which has been investigated by another comprehensive body of research, refers to the behavioural process of attributing humanoid qualities to any kind of non-human entity, such as machines (Airenti, 2015; Epley, Waytz, & Cacioppo, 2007; Złotowski, Proudfoot, Yogeeswaran, & Bartneck, 2015). Anthropomorphism occurs across age groups and is already observable in infancy and childhood, when children attribute the unique quality of mindfulness to non-living entities in their environment, such as dolls, robots, or other toys (Airenti, 2015).[5] And although children anthropomorphise such entities according to their fantasy and imagination (e.g. 'as if the doll was living', referred to as 'pretend-play'; Severson & Carlson, 2010), research also suggests that even at early developmental stages children are able to clearly distinguish their fantasies and imaginations from reality (e.g. Samuels & Taylor, 1994; Woolley, 1997). In addition, anthropomorphism – both in case of children and adults – needs sufficient "space for projection" (Airenti, 2015, p.123) in order to allow for anthropomorphism's "creative control" (Severson & Carlson, 2010, p.1101) over the respective entity.

Hence, in the context of the given inquiry, the important question remains what role anthropomorphism could play in the context of DVAs. For instance, how much creative control can be exercised over a DVA? How much space does it offer for imaginative projection? Could a child engage in pretend play (i.e. 'as if the DVA was living')? An important aspect might be indeed that Alexa et al. offer little space for anthropomorphism's creative control mechanisms due to their *explicit* pre-programmed identities and *real* interactive capabilities, therefore limiting the imaginative scope for 'behaving-as-if' engagements and 'pretend play'. However, anthropomorphism also remains an argumentative shortcut from the research perspective on humans' cognitive development, because the attribution of humanoid qualities to non-human entities requires us to have a categorical belief system in order to differentiate what counts as human and what not in the first place. Thus, the third section sheds light on the question how we develop such categorical belief systems that enable us to anthropomorphise things in our environment.

[5] Essentially, this attribution itself is driven by the *perceived*, *believed*, or *pretended* presence of mental states, therefore bestowing the human quality of mindfulness upon the non-human nature of non-living entities (Airenti, 2015; Waytz, Cacioppo, & Epley, 2014). Although anthropomorphism can also occur with respect to living entities, such as animals (e.g. behaving 'as if the dog was speaking'), for the sake of conciseness this discussion only considers the particular case of non-living entities.

1.3 Empirical & Theoretical Aspects of Human Cognitive Development

In general, the research domain of developmental psychology "aims to understand the *history*, *origins*, *and causes* of behavio[u]r […] from birth to biological maturity" (Zelazo, 2013, p.3/4, emphasis in original). In this regard, our *cognitive development* remains a decisive factor, because it links our sensual experience of the present with our conceptualised and theorised experience of the past, therefore giving rise to those mental representations of reality which enable us to manoeuvre through daily reality (Gelman, 2013). These miscellaneous representations with different depths and breadths, so-called 'concepts' or 'schemes', constitute the abstract underpinning for our *emerging* and *constantly improving* understanding of the world, because we as humans are conditioned to maintain a cognitive 'balance' or an 'equilibrium' between the perceived reality and its conceptualised representations in our mind (Huitt & Hummel, 2003).

In the context of the given study, children's so-called '*Theory of Mind*' (ToM) – sometimes also referred to as 'naïve psychology' (cf. Bernstein & Crowley, 2008) – remains of particular importance, because it refers to the developmental stage of recognising other entities as mindful beings (Astington & Hughes, 2013). However, before applying this concept to intelligently behaving machines, such as DVAs, the following section provides a brief theoretical argument why it is even a valid consideration to apply a concept of *mindfulness* to *mindless* machines – a completely overlooked issue in the literature up to this point!

1.3.1 Possible Minds: The Philosophical Nature of Intelligent Machines

The theoretical question whether machines can have mindful brains – or differently spoken, the question whether the inner mechanics of non-living entities in the physical world can possess mentality – has been a philosophical debate ever since René Descartes' (1596–1650) *dualism* on the distinct nature of the *conscious mind* and *unconscious matter* (Wheeler, 2008). However, ever since the emergence of computational MI in the 1950s, mathematicians, engineers, and other MI-pioneers have perceived themselves as "creators of life" (Turkle, 2005, p.237), therefore implicitly taking a *functionalist* philosophical perspective and assuming that our mind is reproducible by modelling the brains' information processes. Thus, it is also not surprising that most of today's world leading MI-pioneers have a background in neuroscience or cognitive psychology (Ramakrishnan, 2019), because, after all, the idea to make machines intelligent by programming self-learning artificial neural networks was inspired by computer scientists' basic belief that the human brain might function in the same way (Hinton, as cited in Thompson, 2019), and that its perfect artificial reproduction remains

entirely possible, at least in theory (Wilczek, 2019). In contrast, modern dualist opponents like Dreyfus (1972, 1992) or Searle (1990, 2009) argue that even the most accurate and sophisticated computational model of the human brain cannot be considered as a thinking and conscious mind in itself. Hence, this philosophical debate comes down to the ancient Cartesian question whether the existence of thought and mindfulness *does* or *does not* depend on the existence of an 'I', and the former group's answer is *no,* while opponents insist on *yes* (Turkle, 2005).

But before this debate even gained momentum, the British mathematician Alan Turing (1912 – 1954) published his seminal paper and proposed that the question whether machines *can actually think* should be replaced by the question whether machines convincingly *act like they can think* from a speculator's perspective (Turing, 1950), which became famously known as the 'Turing Test' (TT).[6] More recently, philosopher Coeckelbergh (2009) has re-emphasised on similarly pragmatic lines how our stance towards intelligent machines should not depend upon the *speculative existence* of an artificial 'I', but rather upon our *actual experience* of an artificial 'P', which refers to machines' intelligent *performance*. Although this behavioural – not to say *pragmatically empirical* – notion of intelligence has received lots of criticism ever since its first manifestation in the TT (cf. Karelis, 1986), most critiques ignore an important logical fallacy, which Moor (1976) summarised as a simple confusion between the claim that the speculator's awareness of the machine's internal working mechanism "*might alter* a justified inductive inference that [it] can think", and the opposing claim that such awareness "*is necessary* to make a justified inductive inference that [it] can think" (p.254, emphases in original). After all, the nature of the first claim is empirical, the nature of the second one remains dogmatic. Moreover, clarifying this confusion also takes some wind out of sceptics' sails who argue that the current degree of manual human labour behind DVAs' functionalities matters: as stated in the introduction, the prompter in the theatre does not necessarily change the audience's perception of the actor's performance, because what counts is the overall magic of the *experience*.

In sum, a minimal yet sufficient conclusion which can be drawn in the context of the given inquiry is that not only potential answers to the question whether machines could be(come) mindful and thinking entities remain ambiguous (i.e. functionalism vs. dualism), but also the relevance of the question itself seems challengeable (cf. Turing, 1950). Therefore, applying

[6] The TT is discussed in more detail below (2.3.1).

ToM to intelligent machines, such as DVAs, cannot be rejected entirely on the grounds of academic arguments – only on the grounds of those philosophical dogmata that believe in the sanctity of human thought, which Turing (1950) himself eloquently anticipated as the "heads-in-the-sand" objection (p.442).[7]

1.3.2 The Application of ToM in the Context of Intelligent Machines

Following the behavioural notion of intelligence, an intelligent machine – *regardless* of its technological sophistication or internal working mechanisms (e.g. remotely controlled vs. entirely MI-driven) – has capacities that make it appear intelligent, therefore potentially conveying the possession of an inherent *mental* status from the speculator's perspective. A foundational developmental pre-requisite for this perception is the gradual attainment of ToM: during this developmental process, children begin to recognise and attribute the qualities of mentality and intelligence to other entities in their environment (e.g. intentionality, cognition, emotionality, or simply consciousness; Astington & Hughes, 2013). This attainment, which normally becomes fully developed during pre- or primary school years, is the result of experience-based *theorisation* and *conceptualisation* that shape an individual's understanding of as well as behaviour towards other living beings, such as humans (Bjorklund, 2013; Wellman, Cross, & Watson, 2001). Or more generally speaking, since theories serve as "maps of reality [that] guide us through unfamiliar territory" (Griffin, 1994, p.4), ToM is our evolutionary determined and behaviourally driven blueprint that enables us to engage with other intelligently behaving and thus potentially mindful beings.

Hence, this argument naturally extends to intelligent machines, because as Saygin, Cicekli, & Akman (2001) point out: "if behavioral evidence is acceptable for granting intelligence to humans, this should be the case for machines as well" (p.483). This also constitutes a fundamental defence of the TT's behavioural argument mentioned above (**1.3.1**), which claims that instead of focusing on the theoretical nature of humans' and machines' internal working mechanisms, we should rather investigate empirically how we engage with entities that seem to behave intelligently (cf. Turing, 1950).

In sum, ToM, a foundational step of children's cognitive development, constitutes an important concept in the context of the given inquiry, because it shapes our understanding of as well as behaviour towards any intelligently behaving entity, including machines. Thus, it

[7] On a scientific side note, as Gopnik (2019) points out in her recent essay, there might even be deep conceptual parallels between the cognitive processes of human brains and machines – parallels that yet remain to be fully discovered by computer science, neuroscience as well as psychology. For a more detailed overview of this argument see the excursus in **Appendix I**.

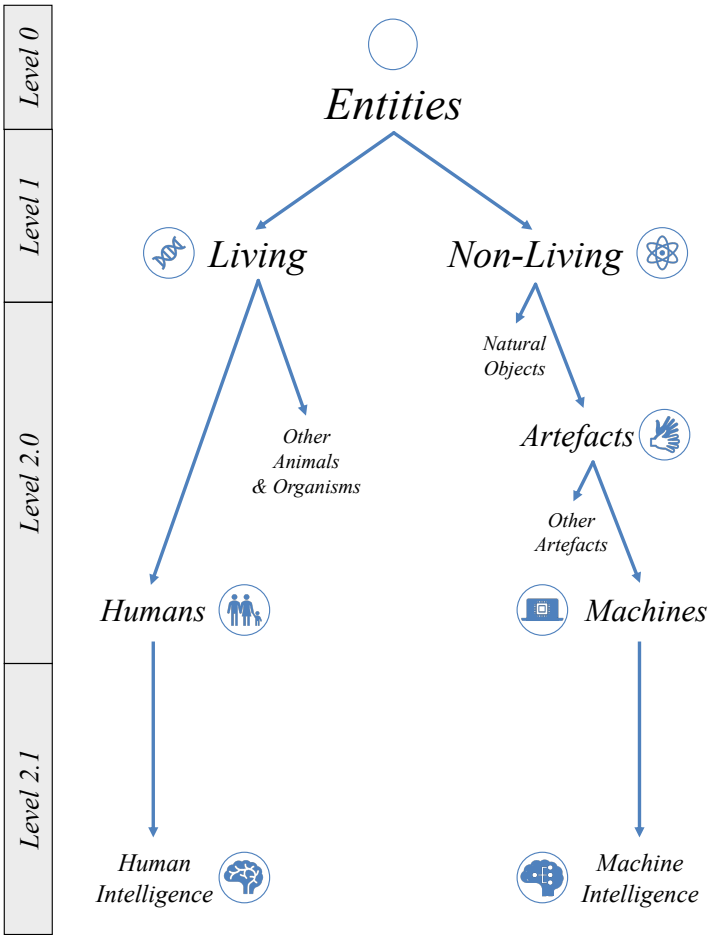
could also be an essential framework in order to understand how children engage with DVAs, because these machines can already convey a strong behavioural notion of intelligence by communicating through voice-user interfaces, answering questions, or fulfilling basic tasks.

1.3.3 Ontological Belief Systems: How We Make Sense of the World

According to the *Piagetian* notion of cognitive development, the entirety of concepts and schemes we develop throughout childhood emerges from a few foundational representations of reality (Gelman, 2013). Hence, from this holistic perspective, the mental framework of such foundational concepts and schemes forms an *ontology* (**Figure 1**), a system of boundaries that allows for a basic categorisation of entities along the lines of their perceived features and attributes, while these boundaries themselves are subject to our innate conditions as well as the sociocultural contexts we are embedded in (Gaudiello et al., 2015).

Thus, different *ontological categories* represent foundational distinctions between the broadest classes of entities (i.e. any subject or object of distinct existence) that we adapt while growing up in order to make sense of the world around us (Severson & Carlson, 2010). Among ontological categories, one of the earliest ontological distinctions that children gradually advance throughout their development is the difference between animate *living* and inanimate *non-living* entities (Wellman & Gelman, 1992). Furthermore, within the living domain, *humans* form a major ontological category in order to differentiate them from other forms of animate life (e.g. animals, plants, other organisms). In turn, within the non-living domain, *artefacts* constitute another major ontological category that comprises any humanly made entity, such as *machines*. Lastly, once children develop ToM, the human ontological category can be extended by the *perceived quality* of HI, and – taking the arguments of the previous discussion into account – intelligently behaving machines can be extended by the *perceived* possession of MI.

Figure 1. Exemplification of an ontological framework



Notes. Figure visualises an exemplification of a multi-levelled ontological framework. *Level 0* refers to the overarching category of entities (i.e. any subject or object of distinct existence). *Level 1* refers to the distinction between the domains of living or non-living entities. Within the domain of non-living entities, *Level 2.0* refers to the distinction between artefacts (i.e. any humanly made entity) or non-living natural objects, and the further distinction between machines (i.e. any technological system) or other artefacts. Within the domain of living entities, *Level 2.0* refers to the distinction between humans or other animals or organisms. Lastly, *Level 2.1* extends the human category and the machine category with the perceived ontological quality of intelligence.

Source. Compiled by SI.

1.3.4 The Ontological Nature of DVAs: Marginality vs. Category in Its Own Right

In the context of the given inquiry, the question of great importance remains to what extent intelligent machines, such as DVAs, may constitute *marginal* entities within children's emerging ontologies, that is, entities without a clear categorical belonging. However, this idea is certainly not new, because even decades ago simple machines were able to emulate particular traits of HI (e.g. robots), while embodying the physical traits of their artefactual nature at the same time (Scaife & Duuren, 1995; Turkle, 2005).

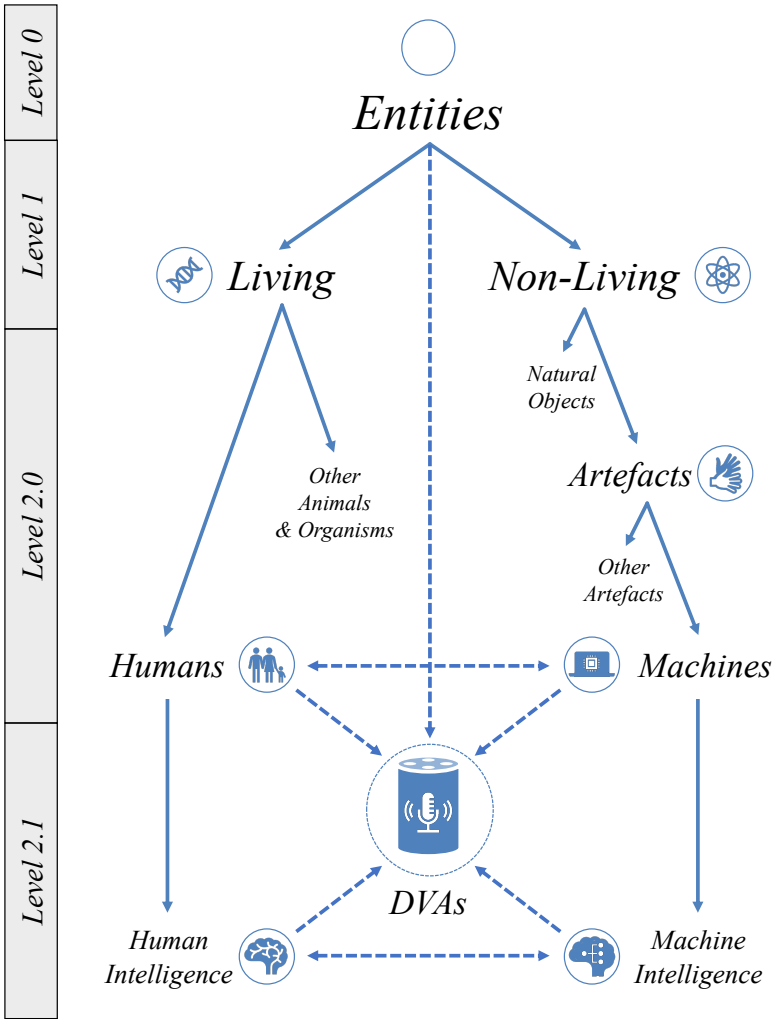
Indeed, previous research has provided ample of evidence for children's unique ontological perception of certain intelligent machines: while strictly distinguishing biological entities as living or non-living beings (i.e. *ontological firmness*) in order to draw further inferences from

these distinctions (Gelman, 1988), children do not hesitate to assume that a robot may have a brain but no heart or to attribute mental states, despite being fully aware of the robot's non-living artefactual nature at the same time (i.e. *ontological openness*; Jipson & Gelman, 2007; Scaife & Duuren, 1995). And although age seems to be an important factor in itself (Cameron et al., 2017; Druga et al., 2017), empirical findings also emphasise the role of parental influences and how they shape children's perceptions of such technologies (Druga, Williams, Park, & Breazeal, 2018; Jipson, Gülgöz, & Gelman, 2016), for instance based on how much technological exposure they allow children to have (Bernstein & Crowley, 2008).

Taking these empirical aspects into account, one can follow Bernstein & Crowley's (2008) conclusion by stating that "children raised increasingly amid intelligent technologies will grow up thinking differently about some concepts that developmental psychologists have previously considered universal and inevitable" (p.242). In light of the technological progress that took place during the last decade since this statement was published, this hypothesis should certainly be considered as an issue of *increased* rather than *decreased* importance.

But then, the question of interest becomes whether some of today's intelligent machines – such as DVAs, which might already characterise childhood environments to a great extent – are *truly marginal* in the sense that they remain entities without a clear categorical belonging from children's ontological perspective. If yes, according to what is known about human cognitive development, one could anticipate that ontological belief systems adjust in order to form a *new* category for such marginal entities, therefore solving cognitive ambiguities and dissonances related to marginality. Or metaphorically speaking, most people would agree that the colour orange is more than a *mere combination* of yellow and red, because we perceive it as a colour *in its own right* (Kahn, Gary, & Shen, 2013). This refers to the basic underlying reasoning of the 'New Ontological Category Hypothesis' (NOCH): while for thousands of years the perceived ontological boundaries of our environment have remained rather stable, only recently technological developments have started to challenge some of the most fundamental certainties in our ontological belief system, such as the superiority of the living domain regarding intelligence in general and HI in particular (Gaudiello et al., 2015; Kahn et al., 2011; Kahn, Kanda, et al., 2012). However, the concrete question whether today's commercial DVAs could constitute an empirical case for the NOCH remains a completely unexplored issue up to this point, despite the fact that – in theory – these concrete humanoid manifestations of intelligent machines could establish themselves on the edge between the living and non-living domain, and thus in the centre of the ontological human-machine framework (visualisation of the NOCH in **Figure 2**).

Figure 2. DVAs' embeddedness in the ontological framework according to the NOCH



Notes. Based on the previous depiction (**Figure 1**), this figure visualises DVAs' theoretical embeddedness in the ontological framework according to the NOCH, meaning that they could constitute an ontological category in their own right. In line with the NOCH, this visualisation places DVAs on the edge of the living and non-living domain, as well as between the ontological categories of humans and machines (i.e. due to the fact that DVAs are particularly humanoid manifestations of MI, for instance because of their humanoid voice).

Source. Compiled by SI.

Furthermore, the claim of the NOCH that intelligent machines, such as DVAs, could constitute an ontological category in their own right is different from the argument of anthropomorphism: as mentioned above (**1.2**), the attribution of humanoid qualities to non-human entities requires us to have a categorical belief system in the first place in order to differentiate what counts as an animate human being and what not. Therefore, the phenomenon of anthropomorphising an inanimate entity (e.g. 'pretend play' or behaving 'as if something was living') occurs *across* existing ontological categories, but it does not *alter* the categories in themselves. In contrast, the argument of the NOCH precedes anthropomorphism: based on children's early exposure to intelligent machines – and parallel

to the development of other ontological categories – a *new* distinct category might emerge simultaneously in children's belief system.

Lastly, the NOCH also challenges the original rationale of the TT to a certain extent (**1.3.1** and **1.3.2**): Turing (1950) followed the basic rationale of an 'imitation game', meaning that someone or something is supposed to convey the *false* notion of being someone or something that is completely different compared to its *true* ontological nature. However, the NOCH would implicitly argue that a DVA could never pass the TT by fooling children into believing that it is human from an ontological perspective, because – according to the NOCH – it would be perceived and *accepted* as an ontological entity in its own right.[8]

1.4 Research Gap & Research Questions

Generally speaking, a relevant research gap opens up if we suspect that what we know as researchers remains insufficient to apply it to the unknown. Up to this point (i.e. spring 2019), the only empirical inquiry dedicated specifically to children's engagement with and perception of commercial DVAs has been conducted by Druga et al. (2017): in their exploratory mixed methods design, the authors gave 26 children (3 to 10 years-old) the opportunity to engage with DVAs (e.g. Alexa or Google Assistant) and a few other interactive toys. Based on interactive questionnaires and more detailed interviews with selected participants, the authors investigated how children perceived DVAs' intelligence, identities and playfulness. In total, their findings showed how children explored and perceived DVAs in various ways while engaging with them; for instance, especially older children (i.e. 6 to 10 years-old) systematically probed how accurately, satisfactorily and consistently DVAs could answer knowledge-based questions, what identity the machines possessed in terms of personal traits or technical qualities and how emotional the engagement with DVAs felt in terms of playfulness or empathy (II in **Table 2**).

[8] On a side note, the basic idea of the NOCH has also been the implicit starting point for a recently published academic manifesto by Rahwan et al. (2019): the interdisciplinary group of researchers proposed that MI should be studied the same way social science research studies human behaviour and development or ethology studies animals. Thus, the authors implicitly assume the validity of the NOCH, because once MI becomes the subject of behavioural and developmental sciences, we have tacitly accepted the respective machines as entities in their own right.

Table 2. Children's engagement with DVAs and observed probing schemes

(I)





(II) Children's quotes related to the exploration of perceived intelligence, identities and playfulness:

Perceived Intelligence:	*Alexa, she knows nothing about sloths, I asked her something about sloths but she didn't answer ... but Google did answer so I think that's a little bit smarter because he knows a little more.*
	[Google Home] *is probably already is as smart as me because we had a few questions back and forth about things that I already know.*
	She doesn't know the answer, ask the other Alexa!
	Alexa what kind of nut am I holding?
Identities:	*What is your favorite color* [Alexa]*?*
	Do you [Alexa] *have a phone inside you?*
	What are you [Google]*?*
Playfulness:	*Hey Alexa, how old are you?*
	Is it OK if I eat you [Google]?

Notes. Part (I) of the table illustrates the research setting used by Druga et al. (2017) in order to allow children to engage with DVAs (left picture: Google Assistant; right picture: Alexa). Part (II) of the table quotes children's comments or observed engagements related to different themes that Druga et al. (2017) reported in their findings.

Source. Druga et al. (2017).

In their concluding remarks, Druga et al. (2017) emphasised how children in their study generally believed that "they could teach the agents, and [that] they could learn from them" (p.599), further implying that DVAs could become a new kind of "learning companion" (p.599) during infancy and childhood. Based on this conclusion, and inspired by the early work of Van Duuren, Dossett, & Robinson (1998), Druga (2018) dedicated her further research to children's agency when programming intelligent machines themselves, and the question how such activities might influence children's perception as well as understanding of respective entities. In contrast, the current study attempts to advance on her previous work (i.e. Druga et al., 2017) by exploring the actual presence of commercial DVAs within today's

natural childhood environments, children's concrete engagement with them, and how these intelligent machines might alter their ontological belief system. In order to guide this exploratory empirical inquiry, four research questions can be formulated (**Table 3**)

Table 3. Overview of research questions

No.	Research question	Type
(I)	To what extent do DVAs constitute a change in today's childhood environments?	*Exploratory*
(II)	What is the nature of children's engagement with DVAs?	*Exploratory*
(III)	How are DVAs embedded in children's general ontological belief system?	*Exploratory*
(IV)	Could DVAs constitute marginal entities or even a new ontological category within children's ontological belief system?	*Exploratory*

Notes. Table shows the four research questions that guided the empirical inquiry. All questions were supposed to build upon each other's empirical essence, and therefore they are also answered within a single integrative discussion in section 3.2 of this paper.

2. Methodology & Research Design

Based on the research questions derived above (1.4), this section outlines the overall methodological approach of the empirical inquiry, including its foundations (2.1), the research design (2.2), the methods of data collection (2.3) and data analysis (2.4), and methodological considerations regarding reflexivity and rigour (2.5) as well as ethics (2.6).

2.1 Methodological Foundations

In general, research paradigms constitute the philosophical foundation of a scientific inquiry and determine the nature of the entire research process, especially regarding methodological considerations (Creswell & Creswell, 2018). These paradigms usually refer to axiomatic *epistemological* and *ontological* assumptions, meaning that they clarify the researcher's basic belief how human knowledge is gained or limited (i.e. epistemology) and what the empirical reality we observe consists of in its essence (i.e. ontology; Martinich & Stroll, 2017; Simons, 2009).[9] While there can be vehement dogmatic tensions between different paradigms, philosophical *pragmatism* follows a strong application-oriented rationale that mostly adheres to the implicit methodological requirements of an inquiry's research questions (Mackenzie & Knipe, 2006; Tashakkori & Teddlie, 1998), therefore making it the most common paradigm of mixed methods research within social and behavioural science (Tashakkori & Teddlie, 2010). From an epistemological and ontological point of view, the widely recognised paradigm of *Deweyan* pragmatism states that "we are participants in an ever evolving, unfinished universe" where inferential knowledge about ourselves and our environment can be acquired – not to say *constructed* – "through *the combination of action and reflection* [emphasis in original]" (Biesta, 2010, para.46).[10]

Since "[m]ethodology is ever the servant of purpose, never the master" (Greene, 2007, p.97), the subsequent empirical enquiry is supposed to follow the pragmatic application-oriented rationale of the previous discussion, therefore pointing in the direction of a *pragmatistic* methodological paradigm that could leverage the research questions derived above (**Table 3** in 1.4) by investigating them through the most accurate methodological lenses. On the one

[9] The use of the term '*ontology*' in this methodological context should not be confused with its previous use in section (1): in a methodological context, ontology refers to a particular belief or claim about the nature of being within a philosophical system of thought – mostly of academic origin – while in a developmental context ontology refers to an individual's cognitive framework for the basic categorisation of perceived entities.

[10] Although a rather narrowed interpretation of Deweyan pragmatism gives a strong preference to interventionalist designs (Biesta, 2010), more open views suggest that pragmatistic methodological considerations should be miscellaneous and eventually yield to an inquiry which is "directly linked to context, to practice, to everyday life, and to lived experience" (Greene & Hall, 2010, para.48).

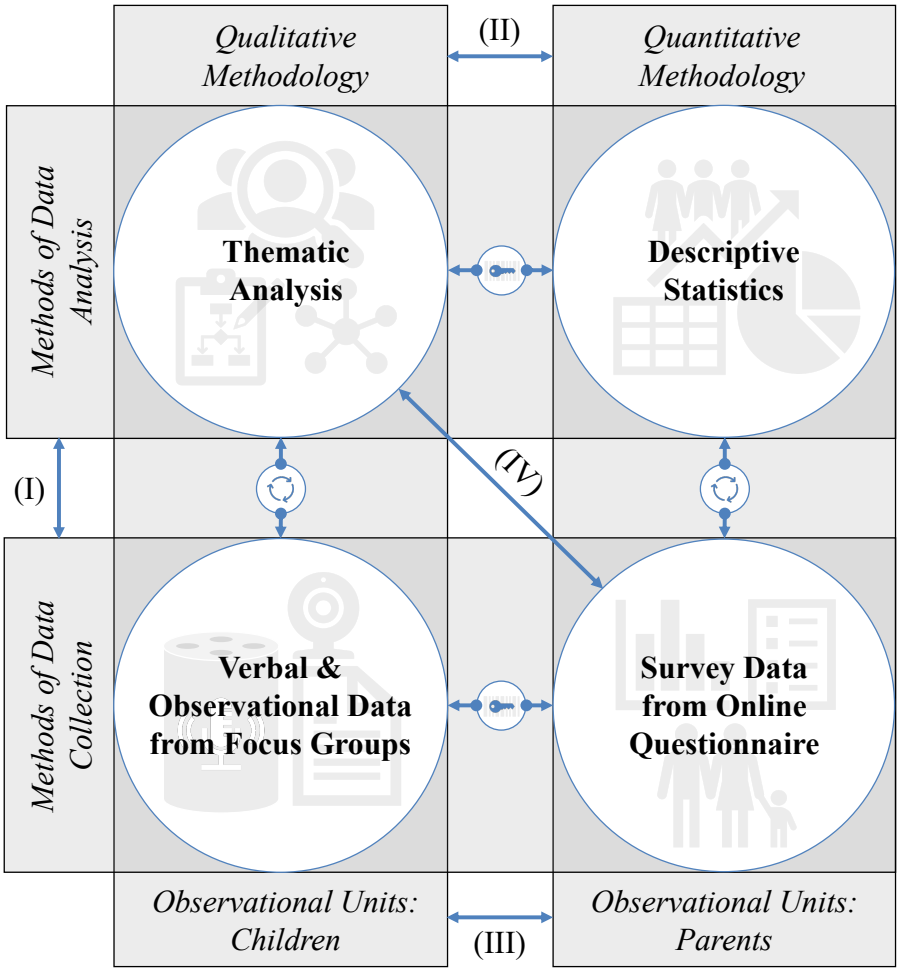
hand, especially the exploration of children's engagement with DVAs as well as their ontological perceptions (II and III in **Table 3**) points in the direction of a qualitative empirical methodology, because instead of reducing the complexity of empirical reality to a few pre-defined variables, qualitative research aims to explore or even understand phenomena as a whole and within their real world context (Flick, 2014). Therefore, whenever the research issues at stake do not allow us to test what is already known in a *deductive* way, the *inductive* approach within the qualitative methodological rationale can help to make sense of new social phenomena and infer meaning from the subjective experience of individuals (Flick, 2014). On the other hand, the promising depth of qualitative insights – sometimes critically referred to as a simple "assembly of anecdote and personal impressions" (Mays & Pope, 1995, p.109) – comes with an array of costs, which are discussed in more detail later (**2.5**). Moreover, investigating a naturally occurring phenomenon, namely to what extent DVAs constitute a change in today's childhood environments (I in **Table 3**), could also require a more structured empirical approach (e.g. survey method). But from a pragmatistic point of view, the systematic combination of different means of scientific inquiry allows for so-called '*triangulation'*, meaning that the strengths and weaknesses of distinct methods or data sources can complement each other by integrating their findings within the same empirical inquiry (Creswell & Creswell, 2018; Denzin, 1970; Jick, 1979).[11] Based on this pragmatistic methodological reasoning, the following discussion specifies a mixed methods research design which was used for the empirical investigation of the research questions.

2.2 Research Design Choice

Generally speaking, a research design is the logical approach chosen and specified by the researcher in order to pin down the means of scientific inquiry based on pre-defined objectives, such as research questions (DeForge, 2010). Essentially, this study followed Druga et al.'s (2017) mixed methods approach by *complementing* as well as *confirming* the qualitative analytical depth of children's unstructured verbal and observational focus group data with the structured breadth of parental survey data collected through an online questionnaire. Hence, in addition to the conventional unidimensional means of scientific inquiry *within* methods and *within* data sources (I in **Figure 3**), this research design incorporated multidimensional triangulation.

[11] Although triangulation between methods or data sources remains the most common application in today's landscape of mixed-methods designs, the concept of triangulation is not necessarily limited to these dimensions, because it can also be used across theoretical perspectives or individual researcher perspectives (cf. Kelle, Kühlberger, & Bernhard, 2019).

Figure 3. Overview of methods and means of triangulation



Notes. Figure visualises the methods and means of triangulation used in the given research design. The vertical linkage (I) and circular arrows indicate the absence of triangulation (i.e. unidimensional investigation *within* methods as well as *within* data sources). The horizontal linkage (II) indicates triangulation *between* different methods of data collection and analysis (i.e. between the qualitative and quantitative methodological paradigm). The other horizontal linkage (III) indicates triangulation *between* data sources (i.e. different types of observational units contributed to the breadth of data collection and analysis). The diagonal linkage (IV) indicates triangulation *within* data sources but *between* methods (i.e. thematic analysis was applied to qualitative survey data). Key code symbols indicate that between subject data triangulation linked data on an individual level, meaning that individual focus group units (i.e. children) were matched with corresponding observational units in the survey (i.e. parents) based on a unique key code.

Source. Compiled by SI.

Firstly, from a methodological perspective, *between* methods triangulation (II in **Figure 3**) links the qualitative means of data collection and analysis (i.e. verbal and observational data from focus groups which were used for thematic analysis) to the quantitative means of data collection and analysis (i.e. descriptively analysed survey data collected from parents through

an online questionnaire).[12] Moreover, triangulation *between* data sources linked the survey insights from parental observational units with data collected from children (III in **Figure 3**).

One of the particular methodological strengths of the given research design was that triangulation between data sources linked data on an individual level, meaning that individual focus group units (i.e. children) could be matched with corresponding observational units in the survey (i.e. parents) based on a unique key code (key code symbols in **Figure 1**; further details in **2.3.2**). Secondly, qualitative text data collected through several open entry items in the parental survey were also used for a thematic analysis, therefore allowing for a triangulation *within* data sources but *between* methods (IV in **Figure 1**).

2.2.1 Sampling Procedure

Following Robinson's (2014) widely-cited sampling guide, this section provides a step-by-step overview of the sampling procedure that was applied.

(1) Target population

When it comes to scientific inquiries that partly rely on children's verbal data, one major methodological challenge is that participants might struggle to reflect upon their own thoughts or actions and verbalise them in order to express the cognitive concepts they already possess (Gelman & Frazier, 2012). Although this issue is certainly not limited to the *pre-operational* stage of children's development, one can assume that during the *concrete operational* stage participants would be able to share deeper and more comprehensive insights into their thought processes, especially due to growing levels of logical reasoning and the acquisition of other fundamental concepts, such as ToM (cf. Gardner, 2011; Huitt & Hummel, 2003). Therefore, in light of the given research questions and the research design, the theoretical *target population* of this inquiry consisted of children who were in the *concrete operational stage* of their cognitive development (i.e. *latent inclusion criteria*).

But the question remains how this population could be identified in a feasible way, because on the one hand there can be tremendous individual differences across children's developmental trajectories (Gelman & Frazier, 2012), but on the other hand an individual a priori assessment of each participant's developmental stage in order to determine eligibility would have exceeded the inquiry's limited resources. Hence, for the pragmatic sake of

[12] For the sake of conciseness, this part of the discussion narrowly assumes that survey methods are part of the quantitative methodological paradigm. The short-sightedness of this assumption is considered in more detail below (**2.3.2**).

practicability, this scientific inquiry relied on an age range between 6 and 10 years (i.e. *manifest inclusion indicator*), because from an empirical perspective the majority of normally developing children reaches the concrete operational stage at some point within this period (Huitt & Hummel, 2003). Since this particular age range roughly corresponds to the period of primary school education in most industrialised countries (OECD, 2014), the sampling strategy of the given inquiry was based on co-operations with primary school institutions, which offers various methodological as well as practical advantages (cf. Rice, Bunker, Kang, Howell, & Weaver, 2007; further discussed in **2.5**). In sum, the target population of this inquiry was defined as normally developing primary school children between the age of 6 and 10 in industrialised countries.

(2) Sample size and sampling strategy

In light of the inquiry's limited resources, determining the sample size as well as the sampling strategy also constituted a pragmatic compromise between methodological as well as practical considerations. At its very core, the purpose of this inquiry was exploratory, because it was supposed to investigate a new phenomenon and to derive *possible* explanations of its nature. In general, such "demonstration[s] of possibility" (Robinson, 2014, p.30) are less about the representativeness or generalisability of finite findings and more about the awareness of potent potentialities that become manifest in the data. Therefore, a *case study* constituted a practical sample size framework for this purpose (cf. Robinson, 2014), because it allows for an in-depth analysis of a single case (n=1) that can still have relevance beyond its own context (further discussed in **2.5**). For this inquiry, the selected case was the Student Investigator's (SI) home town, a suburban municipality in Northern Germany called *Chandler*.[13] Hence, the case selection followed the non-random rationale of a *convenience sampling strategy*, meaning that the case was selected from the target population based on the practical criteria of easy accessibility and availability (Etikan, Musa, & Alkassim, 2016; Robinson, 2014). The main arguments of convenience for this particular selection were personal network connections to local primary schools as well as affordable resource expenditures for travel and accommodation costs.

Chandler has approximately 15,000 inhabitants who live in the town Chandler itself or in one of the seven villages that are part of the municipality. Regarding most demographics (e.g.

[13] For the sake of full-anonymity, all names mentioned throughout the discussion are pseudonyms with no similarity to the real names of individuals, institutions or the municipality itself.

age distribution, percentage of immigrants, gender ratio), Chandler lies within ¼ of a Standard Deviation (SD) of the German national average, while most socioeconomic characteristics (e.g. median income, unemployment rates, percentage of social security recipients) slightly exceed ¼ of a SD, therefore indicating that on average rather stable and wealthy socioeconomic circumstances prevail in the population. Chandler's public primary education consists of two schools with a total number of approximately 600 students (school 1 ∽400, school 2 ∽200).

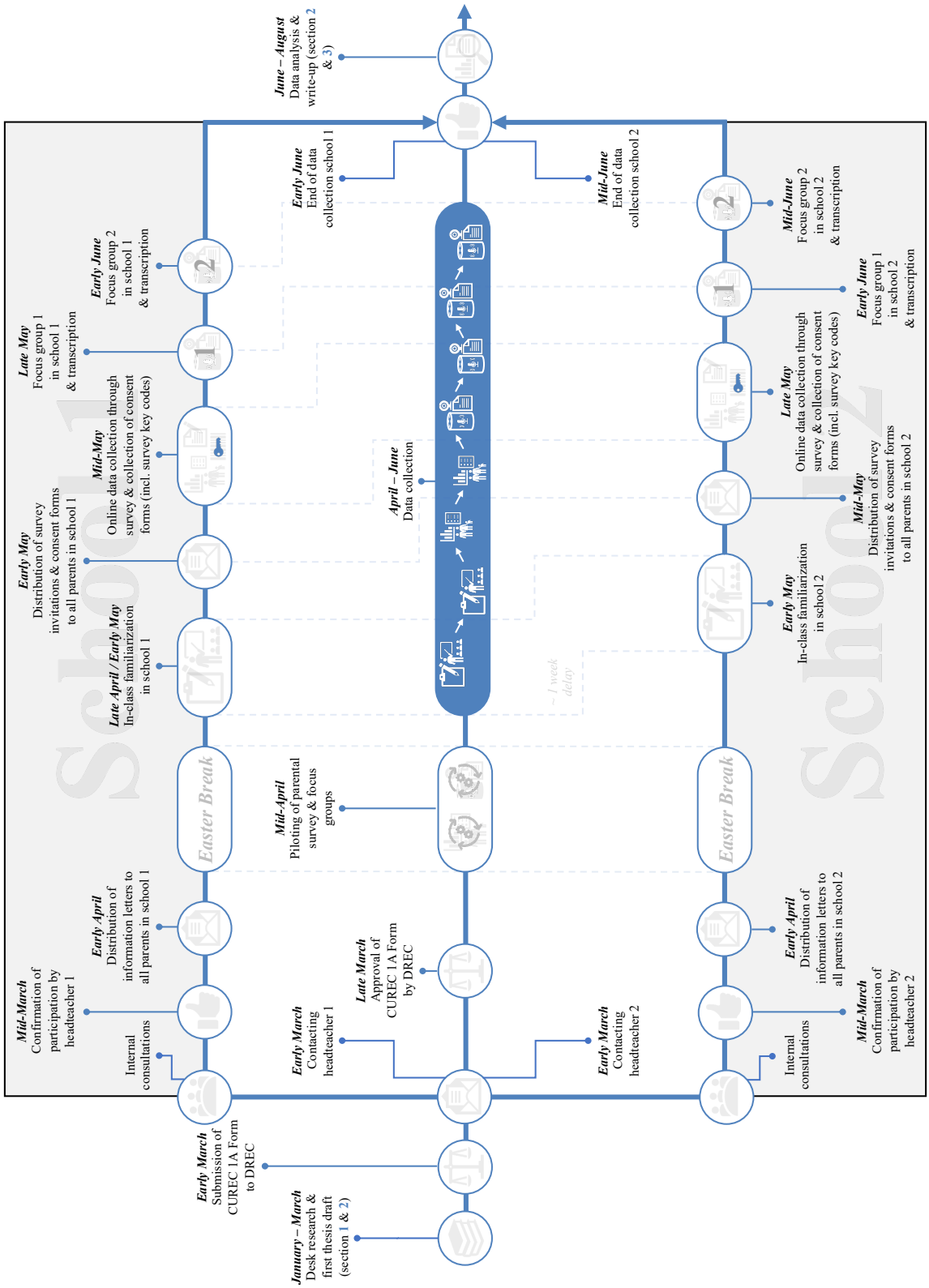
In order to obtain comprehensive in-depth insights into this population of primary school students based on the mixed methods design outlined above, two additional sample size ranges were determined. Firstly, the sample size for the parental survey was not limited by any resource restrictions, and therefore it was supposed to reach a maximum ($n \rightarrow 600$) by distributing survey invitations in the entire population. Secondly, the sample size for focus groups was restricted due to the fact that only the SI himself with the assistance of one volunteering teacher was supposed to conduct the sessions at a given time point, which limited the number of interview groups per primary school to two.[14] Since from a methodological perspective focus groups can contain between $6 < n < 8$ individuals in scientific contexts (Krueger & Casey, 2000) – although slightly lower boundaries (i.e. $5 < n < 7$) should be considered for focus groups with children (Morgan, Gibbs, Maxwell, & Britten, 2002) – the sample size range for the focus groups was between 10 and 14 participants per school (i.e. [2 groups × 5 children] < n < [2 groups × 7 children]). For the selection of focus group participants, *convenience sampling strategies* were applied based on parents' as well as children's willingness to participate and – in case of exceeding willingness to participate – on a practical first-come-first-served selection (cf. Robinson, 2014).

2.2.2 Research Procedure

After the initial desk research was completed in March 2019, the data collection procedure lasted for almost 4 months from early March 2019 to mid-June 2019, including negotiations with the two participating schools, piloting of data collection methods, in-class familiarisation, and the actual collection of primary data through parental surveys and focus groups (visualisation in **Figure 4**, detailed description in **Appendix II**).

[14] According to the original research design, each focus group was supposed to be further divided into *two* interview groups, and both groups were supposed to be moderated separately by the SI himself on the one hand and a volunteering teacher on the other hand (2.2.2).

Figure 4. Chronological overview of research procedure



Notes. Figure visualises the chronological order of the research procedure. This overview depicts all major steps related to the collection of primary data in both schools. All time points refer to the calendar year 2019.

Source. Compiled by SI.

In total, two major complications occurred throughout the data collection: firstly, during the second focus group in school 1, the recording device suddenly stopped working 10 minutes before the session ended. The SI noticed this technical failure approximately two hours later and immediately tried to reconstruct the focus group discussion. Unfortunately, the data for this missing time frame remained very fragmentary. Secondly, right before the first focus group session in school 1 started, the volunteering teacher, who was supposed to assist the SI, dropped out unexpectedly. Therefore, the group of children could not be split in half and the SI made the ad hoc decision to conduct a single focus group interview with all children. Although the resulting number of focus group participants exceeded the desired group size (2.2.1), the first focus group session in school 2 was also conducted with the combined group of children in order to ensure comparability in the data collection procedure.

2.3 Methods of Data Collection

As pointed out above (2.2), *two* methods of data collection were used: verbal and observational data from focus groups constituted the main pillar of this inquiry, complemented by parental survey data. For both methods, this section discusses underlying methodological considerations as well as the applied means of data collection.

2.3.1 Collection of Verbal & Observational Data through Focus Groups

Due to the exploratory purpose of this inquiry, two consecutive focus group sessions were conducted with each group of children in order to maximise the breadth and depth of empirical insights. Although having several focus group sessions with the same group of participants is not entirely unproblematic (e.g. attrition due to participant dropout after the first session), the unique advantage is that a larger variety of aspects can be explored in more detail and with different means of scientific inquiry, which was considered to be more important in the exploratory context of the given inquiry.

(1) Methodological considerations

In general, the collection of verbal data through interview methods relies on the nature of human speech to express individual perceptions, experiences or opinions through the words of a language (Flick, 2014). Therefore, the collection of verbal data *always* depends on interviewees' ability and willingness to express themselves, which is not an issue limited to children, but obviously of particular importance in this context depending on age or individual competence and confidence levels (Eiser & Twamley, 1999). In line with the definition of the target population derived above (2.2.1), the beginning of middle childhood

(i.e. around 6 years) is normally regarded as a feasible age threshold, although the particular nature of an interview as well as children's individual competence or confidence levels can allow for many exceptions from this benchmark (Bricher, 1999; Prior, 2016). From a researcher's perspective, a major challenge remains to translate the complex and abstract issues of interest into tangible and illustrative expressions, while avoiding over-simplistic or naïve notions at the same time (Prior, 2016). In order to keep younger children engaged and focused, structured activities and attentive stimuli (e.g. drawing exercises, playing games, showing tangible objects) are important constituents of an interview design (Mauthner, 1997). In addition, one has to keep in mind that children are often strongly habituated with respect to certain question and answering behaviours, which can be substantially different from the nature of research interviews (e.g. right/wrong and true/false questions, or felt obligation to provide answers at all times; Bricher, 1999; Prior, 2016). The natural power imbalance between the interviewer and the child remains another issue that may be taken into account by integrating empowering activities and allowing participants to remain flexible regarding the pace or even the agenda of the interview (Flewitt, 2014; Mauthner, 1997; Mayall, 1994). Lastly, challenges related to reflexivity arise, because an adult interviewer is naturally constrained by the existing age gap or other potential factors, such as socioeconomic gaps, therefore imposing a perspectivist asymmetry on the interview situation that can influence the research significantly (Kortesluoma, Hentinen, & Nikkonen, 2003).

The use of focus groups, a general sub-category of interview methods (Flick, 2014), requires additional considerations, which are also not necessarily limited to children, but probably of particular importance in the context of young participants. For instance, compared to bilateral one-on-one interviews, the voices of children with rather shy and anxious personalities might be easily drown out by more dominant participants (Hennessy & Heary, 2005). Hence, the moderator should pay special attention to the voices of more quiet and hesitant children (Hannay, Dudley, Milan, Kellogg-Leibovitz, & Rodino, 2016). In contrast, focus groups offer several methodological advantages: besides their efficiency to collect verbal as well as observational data from a larger number of individuals within a short period of time, they also have an inherent quality control due to social checks and balances, which can restrain extreme views or behaviours and enable the moderator to facilitate a common ground among participants (Hannay et al., 2016). Since group settings can resemble peer-to-peer environments of children's daily conversational experiences (e.g. classroom or playground conversations), one could also argue that the natural power imbalance between researchers

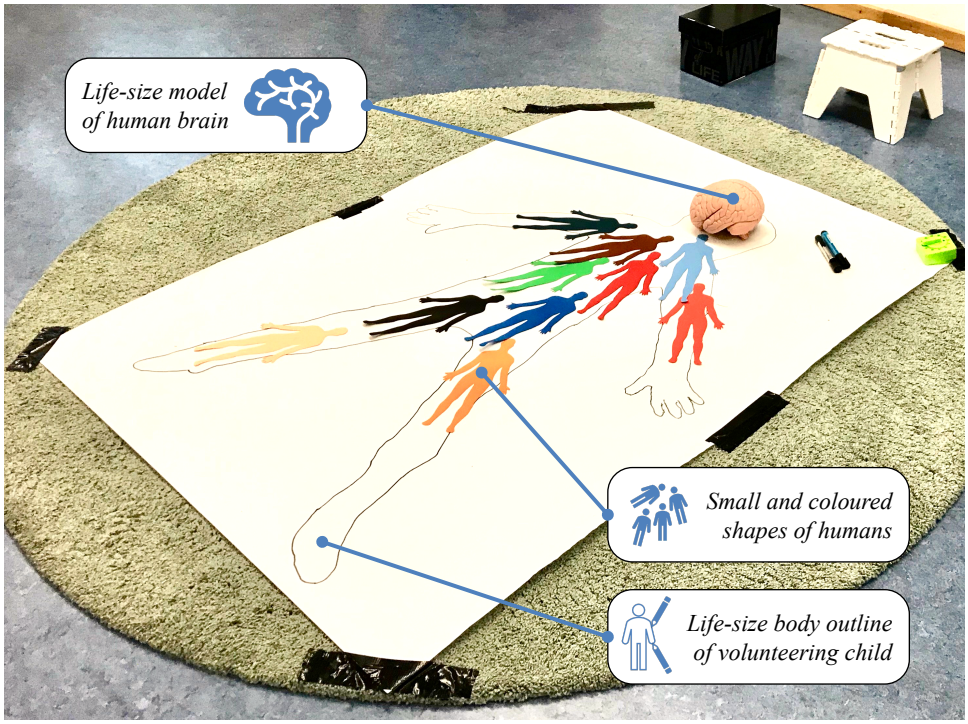
and interviewees becomes weaker and children feel more secure in a group setting, especially if groups are recruited from the same social context (Mauthner, 1997).

(2) Design of focus group session 1

The first focus group session was supposed to elicit and probe children's reflections upon human nature in general and HI in particular, and how they related these reflections to machines in general and MI in particular (detailed focus group procedure in **Table 4**, focus group set-up in **Figure 5**).

The pilot of this focus group session mainly improved the wording of questions and added additional probing schemes. For instance, children's reflections upon human nature turned out to be quite anatomical, and therefore a separate section on HI was integrated with its own stimulus (i.e. life-size model of human brain) to direct the discussion towards more intelligence-based peculiarities of human nature. Moreover, the pilot suggested that note taking during the focus group was not advisable, because it turned out to be more distracting than helpful for children when they tried to decipher the SI's muddled handwriting. Lastly, piloting also confirmed the importance of integrating a break during the focus group, because average attention spans dramatically dropped after approximately 30 minutes of discussion.

Figure 5. Final set-up of focus group session 1



Notes. Figure shows basic set-up of the first focus group session. The picture was taken in the aftermath of the first session in school 1.

Source. Picture taken by the SI.

Table 4. Detailed procedure of focus group session 1

Part	Title	Length	Description
(I)	Introductory briefing	∽ 5	(-)
(II)	Structured activity	∽ 5	Group drawing exercise: • Aim: Engaging children in abstract and categorical ways of thinking • Volunteering child was asked to lie down on large piece of paper; two other volunteers were asked to draw the first child's body outline • Body outline was supposed to serve as first level of abstraction (i.e. volunteering child → volunteering child's life-size body outline) • Following this SI asked the group to think of potential differences and similarities between volunteering child and themselves • For each comment a small and coloured shape of a human was placed in the life-sized body outline in order to symbolize further levels of abstraction
(III)	Moderated group discussion	∽ 15	Questions to probe the discussion: • What does it mean to be a human? / What is special about humans? • What are the similarities and differences among humans? • How do humans differ from other living beings? • What does it mean to 'live'?
(IV)	Structured activity	∽ 5	Attentive stimulus: • Aim: Guiding group discussion towards human brain and qualities of HI • SI asked group what the most important part of the human body might be • After at least one child named human brain SI showed a life-size model of the human brain and passed it around
(V)	Moderated group discussion	∽ 15	Questions to probe the discussion: • Why do we need our brain? / What can we do with our brain? • What is special about our brain? • How do we learn things? / How do we acquire knowledge? • What makes a human very bright and smart?
(VI)	Break	∽ 10	(-)
(VII)	Moderated group discussion	∽ 15	Questions to probe the discussion: • What are the differences between a human brain and a computer? • What are the things that a computer can do better than a human? / Can a computer be smarter than a human? • What are the things that a human can do better than a computer? • Who has heard of artificial intelligence? / What does artificial intelligence mean?
(VIII)	Debriefing	∽ 5	(-)

Notes. Table shows detailed focus group procedure of the first session. The approximate (∽) length of each part is indicated in minutes. Questions were introduced and explained one-by-one, mostly in the depicted order depending on the flow of the discussion.

Source. Compiled by SI.

(2) Design focus group session 2

Based on an interactive game with two commercial DVAs as well as group discussions, the second focus group session was supposed to investigate children's engagement patterns with DVAs. But even more importantly, it was also supposed to elicit and probe children's perceptions of DVAs' humanoid characteristics on the one hand and non-humanoid machine characteristics on the other hand (detailed focus group procedure in **Table 5**). For this reason, the session was based on the 'Human-Centred Turing Test Game' (HCTTG), an activity fully developed by the SI for the purpose of this particular focus group session. Although the HCTTG is based on the original concept of the TT (**1.3.1** and **1.3.2**), several conceptual as well as practical adjustments were made which are discussed in the following.

Since its first publication, the original test concept developed by Turing (1950) has been subject to controversial academic discussions, both within as well as across disciplines (Saygin et al., 2001). Although as a result of these academic discussions many other versions of the TT have been developed or even applied throughout the decades (e.g. Harnad, 1991; Watt, 1996; or the annual 'Loebner Prize', AISB, 2019), their ultimate purpose has always been *machine-centred*, meaning that the objective has been to assess, judge, or even validate the computational 'intelligence' of the machines used in the tests (Saygin et al., 2001). In contrast, the HCTTG is meant to be *human-centred*, meaning that it uses the basic concept of the TT as a *probing tool* to elicit genuine perceptions of machines' humanoid characteristics on the one hand and non-humanoid characteristics on the other hand.

The concrete HCTTG set-up is the following: in line with Turing's (1950) original test (I in **Figure 6**), the basic idea remains that a human subject (i.e. *interrogating party*) can engage bilaterally with two entities of distinct nature (i.e. *machine* and *human*), but through identical language-based communication interfaces (e.g. screen-based text conversations) in order to constrain information regarding each entity's true nature. The task of the interrogating party is to reveal the true ontological nature of entity 1 and 2 based on bilateral communication experiences. The first novelty of the HCTTG is that the aforementioned methodological reasoning of a focus group is applied, meaning that the interrogating party consists of a group of children instead of a single individual (II in **Figure 6**). Secondly, both entities are commercially available DVAs with their off-the-shelf voice-user interfaces. Thirdly, in order to constrain information regarding each entity's true nature, the interrogating party is told that there is a human who *can* remotely control DVA 1 *and/or* DVA 2.

Table 5. Detailed procedure of focus group session 2

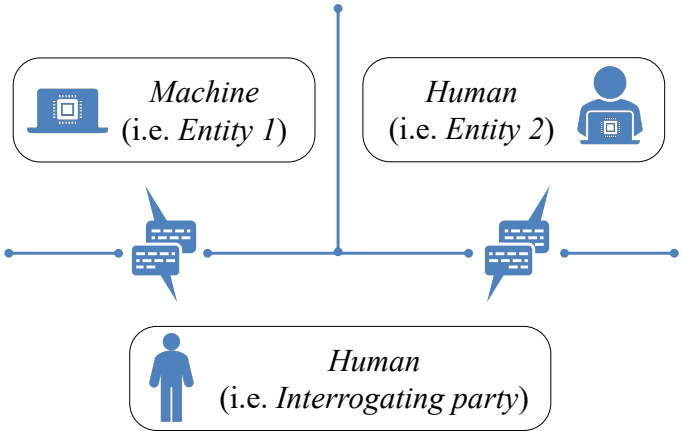
Part	Title	Length	Description
(I)	Introductory briefing	∽ 5	(-)
(II)	Group reflection	∽ 5	Reflection exercise: • Repetition of the previous focus group session mainly based on what children were able to memorise (input was complemented by SI if necessary)
(III)	Introduction to the HCTTG	∽ 10	Introducing the group to the HCTTG: • Presenting DVAs to the group: Alexa (device: 'Amazon Echo Dot') and Google Assistant (device: 'Google Home Mini') • Brief discussion on DVAs in general (e.g. Who has seen a DVA before? Who has a DVA at home? What is a DVA?) • Instructing the group with the rules and objectives of the HCTTG (e.g. emphasising that all means of verbal inquiry are allowed and that reasonings behind answers are as important as the answers themselves) • Separating group into grade-specific groups (i.e. Group 1: first and second grade; Group 2: third and fourth grade)
(IV)	First round of HCTTG	∽ 15	• Group 1 (younger children) in Room 1 (Alexa) • Group 2 (older children) in Room 2 (Google Assistant)
(V)	Break	∽ 10	(-)
(VI)	Second round of HCTTG	∽ 15	• Group 1 (younger children) in Room 2 (Google Assistant) • Group 2 (older children) in Room 1 (Alexa)
(VII)	Moderated group discussion	∽ 15	Questions to probe the discussion: • What DVA could have been controlled by a human? • Why or why not?
(VIII)	Debriefing	∽ 5	Important details of the debriefing • SI not revealed that both DVAs were not controlled by a human being and that the alleged human controller was indeed fictional • SI explained the idea of the original TT as well as the reasons for playing the HCTTG

Notes. Table shows detailed focus group procedure of the second session. The approximate (∽) length of each part is indicated in minutes. During the first and second round of the HCTTG, the SI entered each room once or twice for approximately one or two minutes in order to ensure that children maintained a focus on the task. Both rooms were audio-visually recorded.

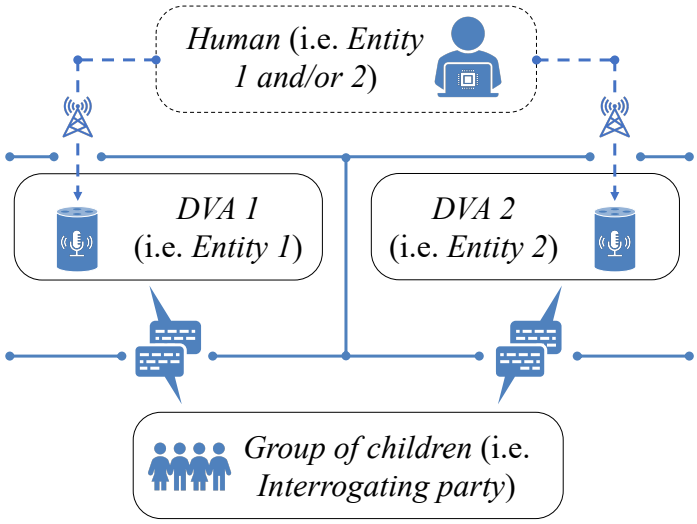
Source. Compiled by SI.

Figure 6. Design details of the TT and the HCTTG

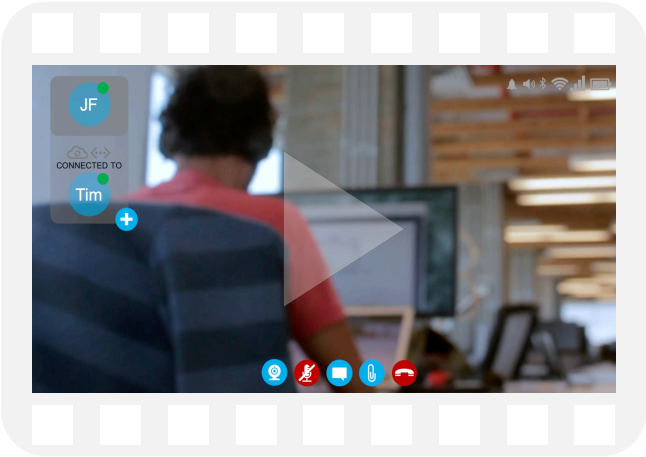
(I)



(II)



(III)



Notes. Part (I) of the table shows the basic set-up the original TT, where the interrogating party engages bilaterally with two entities of distinct nature (i.e. *machine* or *human*), but through identical language-based communication interfaces (e.g. screen-based text conversations). The task of the interrogating party is to correctly identify the true nature of entity 1 and 2 based on bilateral communication experiences. In contrast, part (II) shows the set-up of the HCTTG developed by the SI: the interrogating party consists of a group of children who engages with commercially available DVAs (i.e. entity 1 and 2) through voice-user interfaces. In this set-up, the interrogating party is told that there is a human who *can* remotely control DVA 1 *and/or* DVA 2. Part (III) shows a thumbnail from the video footage that was used in order to introduce interrogating parties to an alleged human controller named 'Tim'.

Sources. (I) adapted from Saygin et al. (2001), (II) – (III) compiled by SI.

But in contrast to other TT-designs, the purpose of the HCTTG is *neither* to examine whether DVAs can pass the test *nor* to assess the interrogators' ability to correctly distinguish DVAs from humans, but rather to elicit what perceived characteristics interrogators associate with human nature on the one hand and with intelligent machines, such as DVAs, on the other hand. Therefore, depending on available resources, actual human control can either be *true,* meaning that there is indeed a human controller behind the curtain, or *false*, meaning that there is none. What matters in both scenarios is the genuine *belief* or the *awareness* of the interrogating party that human control over the entities is indeed *possible*.

In the context of the given inquiry, the chosen DVAs were Alexa (device: 'Amazon Echo Dot'), which remains by far the most widely used DVA among German families (Quantcast, 2019), and the Google Assistant (device: 'Google Home Mini'), which is also one of the most popular DVAs in the German population according to recent market research (Postbank, 2019). In addition, since research resources were constraint, the implementation of actual human control was not feasible. Instead, the interrogating parties were shown footage of a video call with an alleged human controller named 'Tim' (III in **Figure 6**) and told that he would be able to remotely control the Google Assistant (i.e. DVA 1) *and/or* Alexa (i.e. DVA 2). The alleged human controller was depicted while working in front of computer screens and only shown from behind in a blurry way in order to minimise the influence of distracting visual details. After the SI had introduced the alleged human controller, the group of children was asked to explore and find out *whether* he might control one *or* both of the devices and *if so*, *which one* and *why*.

In fact, the pilot of this focus group session yielded to the production of the video footage, because in the beginning of the pilot children kept asking the SI why somebody would control Alexa or the Google Assistant, or how this could even work from a technical perspective, therefore not taking the task seriously before it had even started. This suggested that children needed more concrete and tangible 'evidence' in order to actually consider the possibility that there might be a human controller behind the DVA. In addition, the pilot also suggested that at least partial supervision during the 15-minutes of DVA-engagement would be necessary in order to ensure sufficient levels of task focus.

2.3.2 Collection of Parental Survey Data through Online Questionnaire

In the given research design, the collection of parental survey data through an online questionnaire was supposed to complement focus group data by adding structured child-specific background information to the analysis.

(1) Methodological considerations

In general, surveys are a method for structured investigations of naturally occurring phenomena based on systematic and pre-defined means of data collection as well as data analysis (De Vaus, 2008). Although the terms 'survey' and 'questionnaire' are often used interchangeably, one should note that – strictly speaking – a survey refers to a distinct research method in its own right, in which questionnaires simply serve as a possible data collection technique (De Vaus, 2008). In addition, surveys should not be narrowly regarded as a method of quantitative methodological paradigms, because they allow for both quantitative and qualitative means of data collection as well as analysis (De Vaus, 2008). For instance, if data collection is based on questionnaires, structured quantitative as well as unstructured qualitative data can be collected simultaneously through a combination of closed and open entry questions, which remains a widely-used approach, especially within mixed methods research (Bryman, 2006).

But regardless of its exact nature, a major methodological challenge for any structured investigation of naturally occurring phenomena is to translate the latent concepts or constructs of interest, such as personal characteristics, beliefs, knowledge, and attitudes, into measurable indicators or instruments that can be integrated into the structure of a survey design (De Vaus, 2008). While this implies further methodological challenges regarding the *reliability* and *validity* of applied indicators or instruments (Bryman, 2016), researchers must also consider an array of other practical as well as methodological issues when designing their survey tools, such as questionnaires. For instance, the particular format and composition of a questionnaire (e.g. paper-based vs. online, self-administered vs. interviewer-administered, long vs. short, anonymous vs. identifiable) can yield to distortions in the data caused by accessibility issues, response biases due to social desirability and acquiescence, or simply respondent fatigue due to the length or complexity of the instruments used (De Vaus, 2008). Therefore, piloting questionnaires remains an important step in survey-based research in order to identify and improve such issues prior to the data collection (De Vaus, 2008).

(2) Design of the online questionnaire

Just like any other consideration in research, the decision to use an online questionnaire should be driven by the particular nature and context of the inquiry (Sue & Ritter, 2012). However, the advantages of online questionnaires, including methodological as well as practical aspects, clearly prevailed in the given context (**Appendix III** provides checklist

based on Sue & Ritter, 2012). Hence, a questionnaire based on the online experience management platform 'Qualtrics' was programmed, which consisted of three sections with 32 item batteries and a total number of 53 closed items, 4 open entry items, and 5 non-entry items (overview in **Table 6** and detailed structure in **Appendix IV**). The final estimated response-time was approximately 21 minutes.

While some of the instruments used in the survey were either completely adapted from the literature (e.g. four-dimensional instrument for general attitude towards technology and media by Karrer, Glaser, Clemens, & Brude, 2009), or partly inspired by survey methods used in other contexts (e.g. 'Smart Speaker Consumer Adoption Report' by Kinsella & Mutchler, 2019), some instruments were also completely self-developed due to a lack of feasible options in the literature (i.e. general attitude towards MI or children's engagement with technology and media).

In the concluding section, parental respondents were reminded of the focus groups and asked to indicate a non-binding interest regarding their child's participation. If this non-binding interest was positive, respondents were asked to type in a random four-figure key code and to use the same code on the consent form. Respondents were informed that this procedure would allow for an anonymous matching between survey and focus group data, and they were also assured that consent forms would be treated as confidential and destroyed upon the completion of the study.

The pilot of the questionnaire revealed major issues regarding the actual response time, and therefore several item batteries were shortened or dropped completely. In addition, the wording of self-developed instruments was improved, and additional answering options were added to several multiple-choice instruments. Lastly, the mobile-friendly QR code option was complemented by the alternative to access the survey through an anonymous URL, because some participants in the pilot reported a preference to answer such questionnaires on larger computer screens.

Table 6. Overview of online questionnaire used in parental survey

Section	Section headline	Item no.	Category	Conditional	Constructs & Contents
Intro	(-)	1	Non-entry	No	Welcoming
(I)	*Digital Technologies and media in daily life*	2	Non-entry	No	Section introduction and definition of the term 'digital technologies and media'
		3	Closed	No	General usage of digital technologies and media
		4.1 – 4.16	Closed	Yes	Attitude towards digital technologies and media
		5	Closed	Yes	Prevalence of technological trends and developments in daily work environment
		6	Open	Yes	Reasons for refraining from usage of digital technologies and media
(II)	*Is this smart or what? Artificial Intelligence in daily life*	7	Non-entry	No	Section introduction
		8	Closed	No	Familiarity with the term 'artificial intelligence' or 'AI' (synonym to MI)
		9.1 – 9.5	Closed	Yes	Attitude towards 'AI' (synonym to MI)
		10	Closed	No	DVA-ownership
		11 – 15	Closed / Open	Yes	Details of DVA-ownership and usage
		16	Closed	No	DVA-ownership in closer circle of family and friends
		17	Closed	Yes	Child's general DVA-usage in closer circle of family and friends
		18	Closed / Open	Yes	Reasons for refraining from using DVAs
(III)	*Children and digital technologies and media in daily life*	19	Non-entry	No	Section introduction
		20 – 24	Closed / Open	No	Basic demographics
		25	Open	No	Children's usage of digital technologies and media
		26.1 – 26.5	Closed	No	Attitude towards children's engagement with digital technologies and media
		27	Open	Yes	Child's DVA-usage at home
		28	Open	Yes	Child's DVA-usage in closer circle of family and friends
End	(-)	29	Non-entry	No	Thanking respondents for participation
		30 – 32	Open / Closed	Yes	Non-binding interest in focus groups, and random four-figure key code

Notes. Table shows overview of online questionnaire used in parental survey, including section structure, section headlines, respective item numbers, item categories, whether the items were conditional (e.g. contingency or follow-up questions) and underlying item constructs or question contents. For a detailed overview of the questionnaire see **Appendix IV**. For the questionnaire, the more popular term 'Artificial Intelligence' (AI) was preferred over the term 'Machine Intelligence' (MI) that is primarily used in the context of this discussion.

Sources. Items 4.1 – 4.16 adapted from Karrer, Glaser, Clemens, & Bruder (2009), rest compiled by SI.

2.4 Methods of Data Analysis

In the context of the given research design, the use of mixed methods for data collection (2.3) extended to the means of data analysis, which incorporated the triangulation between quantitative as well as qualitative analytical techniques in order to fully integrate the findings from different data sources within the same inquiry (cf. Onwuegbuzie & Combs, 2010). Thus, the analysis relied on *two* distinct methods of data analysis, namely a qualitative *thematic analysis* (2.4.1) and a quantitative *descriptive analysis* (2.4.2).

As pointed out earlier (2.2), a particular methodological strength of the design was that findings from different data sources were integrated on an individual level, meaning that individual focus group units were matched with the corresponding survey units based on a unique key code. In order to facilitate this particular type of mixed methods integration through computer-assisted means of data analysis, the given study relied on two software tools, namely *SPSS 25* for quantitative analysis and *NVivo 12* for qualitative analysis as well as the integration of findings (cf. Bazeley, 2010; Bazeley & Jackson, 2014).

2.4.1 Qualitative Method: Thematic Analysis

In its broadest sense, qualitative data analysis can be defined as a systematic investigation of unstructured material by "reducing the data into themes through a process of coding and condensing" (Creswell & Poth, 2018, p.183). Thus, as a result of this inferential process an understanding of the data can be obtained, and the researcher can make abstract claims about reality as it is embodied by the observed data. In their widely-cited paper, Braun & Clarke (2006) consider *thematic analysis* as one of the most foundational and flexible methods of qualitative inference, especially in the realm of psychology.

(1) Methodological considerations

At its very core, thematic analysis simply translates into the searching for repeated patterns of meaning – so-called *themes* – in the data (Braun & Clarke, 2006). From a practical perspective, this searching relies on the *coding* of data, which refers to the process of "developing categories, properties, and relations between them" (Flick, 2014, p.403). This process can either follow a data-driven *inductive* rationale (i.e. in case of grounded theory, data yield to themes and themes yield to theory; Glaser, 2002) or a theory-driven *deductive* rationale (i.e. theory yields to themes, and themes are applied to data; Braun & Clarke, 2006). Since both rationales produce findings that are usually very different in their nature (e.g. inductive approaches allow for representative thematic description of data, while deductive

approaches can guide in-depth investigations of existing concepts), their choice should be informed by the nature of the research questions (Braun & Clarke, 2006).

In contrast to Braun & Clarke's (2006) view that thematic analysis constitutes a flexible qualitative method in its own right, others have argued that it simply remains an analytical *tool* which is used across various applications of qualitative research (e.g. Boyatzis, 1998). Similarly, Flick (2014) critically points out that using thematic analysis as a method in its own right comes at the risk of reducing qualitative inquiries to "hands-on procedure[s]" (p.423). However, Braun & Clarke (2006) themselves admit that the methodological nature of thematic analysis would also allow for the more colloquial label "grounded theory 'lite'", since it does not require researchers to "fully subscribe to the theoretical commitments of a 'full-fat' grounded theory" (p.81). This stands in stark contrast to more dogmatic understandings of grounded theory as a "full package decision" (Holton, 2007, p.284) that should guide the course of the entire research process. But Braun & Clarke (2006) challenge such dogmatic attachments with certain methods – which others have labelled as a "method for method's sake" (Holloway & Todres, 2003, p.356), or, even more provocatively, as "methodolatry" (Janesick, 1998, p.48) – and argue that thematic analysis does remain an analytical method in its own right. In particular, the authors argue that it offers a *foundational*, *flexible*, but nevertheless *consistent* route for qualitative analysis without ending up on the one-way road of dogmatic research paradigms (Braun & Clarke, 2006).

(2) Application of thematic analysis

In the context of the given mixed methods design, which itself emerges from a pragmatic paradigm, the implicit flexibility and undogmatic openness of thematic analysis were convincing arguments to apply its lenses to the rich and unstructured focus group data. More specifically, an exploratory and inductive rationale was applied, meaning that the data themselves were supposed to yield to themes, and themes were supposed to yield to a conceptual understanding of the data based on abstract claims about the observed reality. Regarding the assessment of a theme's relevance and prevalence, the analysis also took an exploratory approach by considering its frequency as well as its overall depth and importance in light of the guiding research questions.

2.4.2 Quantitative Method: Descriptive Statistics

One of the most foundational issues of quantitative analysis is the question whether statistical inference can be applied to the data, meaning that beyond numerical descriptions (i.e.

descriptive statistics, such as measures of central tendency or variability) the nature of the sample allows for generalisations based on probabilistic claims (i.e. hypothesis testing) that refer back to the sample's underlying population (Onwuegbuzie & Combs, 2010). But as Leech & Onwuegbuzie (2002) noted: "[d]isturbingly, the majority of studies in the social and behavioral sciences do not utilize random samples" (p.7), despite the fact that this type of probabilistic sampling remains the most fundamental assumption of statistical inference and resulting probabilistic claims (Chow, 2010). But in the context of the given study, the claim that random sampling could be assumed would simply be a *blatant lie*! As pointed out above (**2.2.1**), the sampling procedure followed completely different *non-random* rationales. This should already raise a red flag regarding the applicability of inferential statistics. One could further argue that even in the absence of 'true' random sampling the data could still be approximately representative of the particular primary school population *if* the data set was very large (i.e. $n \rightarrow 600$). However, the final survey sample included less than 12% (n=68) of Chandler's primary school population. Although this could constitute an acceptable percentage share for genuine random samples, the nature of this sample was entirely based on parents' willingness and interest to respond, and therefore it would be highly questionable to assume that the 12% of Chandler's primary school population who chose to respond were representative of the 88% who did not.[15]

Therefore, this inquiry's quantitative analysis relied on non-probabilistic descriptive statistics, such as measures of central tendency or variability (e.g. frequencies, mean and median values, variances and SDs) in order to provide a detailed account of the data, which in turn fed into the qualitative analysis, especially by linking insights on an individual level. Lastly, following the basic principles of thematic analysis outlined above (**2.4.1**), unstructured data collected through open entry items were analysed with a coding procedure in order to condense the data into meaningful themes.

2.5 Reflexivity & Rigour (1): The *Design Quality* Dimension

In its broadest sense, and (almost) regardless of disciplines or methodological boundaries, the idea of academic *rigour* refers to the quality of a scientific inquiry based on its methodological integrity and the trustworthiness of its claims (Given, 2008). When it comes to mixed methods inquiries in particular, the challenge remains that very different

[15] In the absence of random sampling, researchers often compare sample parameters to representative population parameters provided by other available sources (e.g. census data) in order to assess the representativeness of the sample with respect to the underlying population and to identify certain biases (e.g. Andreeva et al., 2015). Unfortunately, Chandler's primary schools could not provide enough demographic information that would have allowed for such comparisons.

conceptualisations and terminologies have developed across the methodological landscape (Tobin & Begley, 2004), such as Lincoln & Guba's (1985) widely-cited benchmarks of qualitative research, or Shadish, Cook, & Campbell's (2002) framework of causal inference for experimental and quasi-experimental designs. Therefore, the quality assessment of mixed methods research has always oscillated between the independent application of very different sets of standards, which yielded to the development of more *integrative frameworks* (O'Cathain, 2010), such as the concept of *inference quality* by Tashakkori & Teddlie (2008). This is considered as one of the most foundational frameworks of academic rigour in mixed methods research, because it fully integrates the principles of methodological openness and parity (O'Cathain, 2010). In this framework, Tashakkori & Teddlie (2008) consider two different quality dimensions: *design quality*, which refers to the appropriateness of the chosen means of scientific inquiry with respect to the research questions as well as the accuracy of their implementation, and *interpretive rigour,* which refers to the credibility of the claims that were derived from data. The former dimension is discussed below, while the latter one is discussed in the last section (**3.3**).

In general, design quality can be assessed according to the design's suitability, fidelity, consistency and its analytical adequacy (Tashakkori & Teddlie, 2008). For all four criteria, the above discussion was supposed to provide an array of detailed arguments in favour of the different design choices made. For instance, choosing the SI's home town Chandler as a single case for this inquiry (**2.2.1**) did not only enable the SI to maximise his personal presence in the field (i.e. SI was able to leverage available resources in favour of the implementation by spending almost 3 months in the field), but it also helped to gain a more intimate access to institutions as well as participants (e.g. SI could build trust towards head teachers, teaching staff, parents, or children by referring to his own primary school years in Chandler), which can be of particular importance regarding the rigour of qualitative means of scientific inquiry (cf. Lincoln & Guba, 1985). In addition, the chosen means of data collection were supposed to be as open and exploratory as possible; for instance, using the HCTTG as a probing tool (**2.3.1**) did not impose any pre-defined restrictions on children's individual understanding of intelligence, because similar to the original TT it "permit[ed] direct or indirect testing of virtually all of the activities one would count as evidence for thinking" (Moor, 1976, p.251).

However, from a more critical perspective, there are also major methodological issues that must be pointed out: firstly, as mentioned above (**2.4.2**), convenience sampling is particularly prone to self-selection biases, meaning that some groups in the specified target population

might be excluded, while others might be overrepresented (Etikan et al., 2016). Especially if the sampling is based on parents' and children's willingness to participate, the presence of such selection biases should be assumed (cf. Anderman et al., 1995; Shaw, Cross, Thomas, & Zubrick, 2015). Hence, due to the convenience sampling as it was applied in this inquiry, generalisations that go beyond the local context as well as the unique composition of the sample should be treated with high caution (Robinson, 2014). But at the same time, and as stated above (2.2.1), the exploratory purpose of this inquiry limited the importance of representativeness, because the focus was to investigate a new social phenomenon and to derive possible explanations of its nature. Therefore, this inquiry followed the notion of *naturalistic generalisation*, meaning that its claims are *general* to the extent that individual readers can relate to the particular nature of this inquiry and may be able to draw parallels to other contexts (Melrose, 2010; Onwuegbuzie & Combs, 2010).

Regarding the proposed means of data analysis, the integration of an immediate transcription and reflection phase after each focus group session was considered high priority by the SI (2.2.2), because from a methodological point of view this preparational step was not only decisive for the SI's familiarisation with the data (cf. Riessman, 1993), but it also constituted the first integral step of the interpretative analysis itself (cf. Bird, 2005; Lapadat & Lindsay, 1999). However, the quality of transcripts in terms of their accuracy and fidelity to detail also determines the quality of the overall analysis (Braun & Clarke, 2006), further implying that transcripts generated by a single researcher can be a major methodological weaknesses. For this reason, throughout the analysis the SI also referred back to original audio-visual recordings in order to account for potential transcription errors as rigorously as possible.

Lastly, in the context of qualitative research, the investigator himself remains the most important mean of scientific inquiry (Flick, 2014) and therefore a decisive – not to say *subjective* – determinant of the design quality in itself. But instead of flatly denying the methodological role of subjectivity in a rigorous research design, one should reflect upon the nature of the potential errors and biases it creates (Hammersley, 2011). For instance, due to the SI's lack of pedagogic experience at the outset of the field work – or for that matter, even the lack of personal parenthood experience should be worth mentioning (cf. Adler & Adler, 1996) – the research procedure incorporated a comprehensive phase of in-class familiarisation in order to gain sufficient levels of practical insights prior to the data collection. Nevertheless, during the transcription of audio-visual recordings, the SI himself noticed how he could have been more engaging with quiet participants, especially in light of the unexpectedly large group size (2.2.1). At the same time, it is hard to deny that the nature

of the focus group sessions deeply reflected some of the SI's biographical childhood idiosyncrasies: as a child, he always preferred being a quiet observer and appreciated when adults respected the integrity of his passive comfort zone. And although it is also hard to deny that the quality of the focus group sessions could have benefitted from additional practical interview experience, Cohen et al. (2018) raise the crucial question: "is it really possible or *desirable* to set aside one's own biography?" (p.26, emphasis added). Probably not …

2.6 Ethical Considerations

At its very core, a rigorous piece of research must also demonstrate *ethical* integrity (Cohen et al., 2018), meaning that beyond methodological or interpretive considerations (**2.5**) the quality of emerging claims depends on how well scientific norms for conduct were respected, especially regarding fundamental moral principles (Resnik, 2015).

In terms of formal frameworks, the ethical integrity of this inquiry's proposed research design not only relied on the 'Ethical Guidelines for Educational Research' as defined by the British Educational Research Association (2011) but also on the institutional approval that was granted by the Departmental Research Ethics Committee (DREC) on behalf of the Central University Research Ethics Committee (CUREC) at the University of Oxford (**Appendix V**). Institutional approval was granted based on the university's approved procedure for non-invasive research methods with children in institutional or non-institutional settings (i.e. AP25, Version 5.1; CUREC, 2019). The implementation of the research design strictly adhered to this approved procedure as well as other CUREC requirements applicable to this research design.

But despite these formal frameworks, it remains the researcher's responsibility to maintain ethical awareness throughout the entire research process, because in practice "[e]thical decisions are contextually situated – socially, politically, institutionally, culturally, personally" (Cohen et al., 2018, p.111). When it comes to scientific inquiries involving children, an overarching question is whether the inquiry intends to conduct research *on* or rather *with* children, which can already have important implications for the researcher-researched relationship (Mauthner, 1997; O'Dell & Brownlow, 2016). For instance, doing research *with* children can imply that language norms related to adult authority and power should be avoided in order to create a sense of intimacy and proximity (Mauthner, 1997).

Therefore, the SI always introduced himself with his first name and offered children to address him in the second person singular ('Du').[16]

In addition, the SI carefully considered organisational as well as design aspects, including appropriate focus group spaces in which participants could feel safe and comfortable, the close availability of teaching staff at all times, sufficient time slots in the beginning of each session to emphasise the matters of anonymity and confidentiality (e.g. who has access to recordings), or to explain the purpose and the background of the research in a tangible way (cf. Mauthner, 1997; O'Dell & Brownlow, 2016). Lastly, participating children were also perceived as being more vulnerable and prone to harm compared to adult research participants (cf. O'Dell & Brownlow, 2016). For instance, whenever children referred to personal experiences regarding their engagement with DVAs at home, unequal domestic access to modern technologies could have yielded to jealousy or feelings of being deprived. Therefore, during the moderation of the focus groups, the SI prevented conversations from becoming too explicit with respect to such unequal experiences of technological engagement at home.

[16] In the German primary school system it remains very common that children have to address teachers or other adults by their last name and by using the second person pronoun 'Sie' instead of 'Du' (Ehlers, 2009), which has important implications regarding the perceived hierarchy of both communication partners (Winchatz, 2001).

3. Results & Discussion

The third and last main section presents the results of the data analysis (3.1), discusses how these results relate to the initial research questions (3.2) and briefly assesses the overall interpretative rigour of this inquiry's empirical claims (3.3).

3.1 Results of Data Analysis

In total, the data analysis was based on n=68 observational units for the parental survey and n=26 corresponding observational units for focus groups with children (Table 7).[17] Sample sizes were relatively equal across both schools as well as focus group sessions (only the second focus group session in school 2 suffered from an attrition of Δn=5). The demographical composition was relatively equal as well, meaning that the ratios of respondents' demographical characteristics rarely differed by more than 15% across the nested structure of the data (Appendix VI and VII). The same also held true for attitudinal survey instruments: summary statistics indicated that mean values of attitudinal scores differed by less than one half of a SD, while SDs themselves as well as minimum and maximum values were also relatively equal within the sample (Appendix VIII). In sum, these descriptive comparisons suggested that there were no major sampling biases in the nested structure of the data.[18] In addition, statistical diagnostics suggested sufficient levels of internal reliability (Appendix IX) as well as moderate levels of positive inter-instrument correlation (Appendix X) for attitudinal instruments, further implying that all three instruments could be used for analytical purposes in the context of the given inquiry.

[17] Since the use of percentages as statistical measures for frequencies can be critical in the context of very small sample sizes, tables in the appendix and in the discussion always provide absolute numbers as well as percentages.

[18] For instance, if there had been major demographical or attitudinal differences between both schools, the data analysis would have been conducted separately for each institution and the integrative merging of findings would have been treated with high caution due to suspected systematic differences between schools.

Table 7. Overview of sample sizes across the nested structure of the data

	Parental Survey (*n*=68)		Focus Groups (*n*=26)		Session 1 (*n*=23)	Session 2 (*n*=20)
School 1	37 (54.4%)	→	13 (50%)	→	11 (47.8%)	13 (65%)
School 2	31 (45.6%)	→	13 (50%)	→	12 (52.2%)	7 (35.0%)
Total	68 (100%)	→	26 (100%)	→	23 (100%)	20 (100%)

Notes. First column of the table shows survey sample size per school and in total. Second column of the table shows sample sizes of corresponding numbers of children who participated in at least one of the two focus group sessions (i.e. either one *or* both of the focus group sessions). Third and fourth column show exact number of children who participated in the first and second focus group session. For instance, in school 1 there were *n*=37 parents who participated in the survey, and *n*=13 of their children participated in at least one of the focus group sessions (*n*=11 in the first and *n*=13 in the second session). All percentages are calculated per column.

3.1.1 Mixed Methods Survey Results

The survey data analysis allowed for a detailed characterisation of DVAs' general presence as well as the concrete nature of their embeddedness in the technological landscape of childhood environments.

(1) Characterising the presence of DVAs in childhood environments

The first main result of the survey data analysis suggested that it was very important to differentiate between DVA-*ownership* and actual DVA-*usage*: since the majority of today's digital devices, such as smart speakers, smartphones, tablets, or even laptops, allow their users to access DVAs, the survey differentiated between three categories: (1) households that owned devices with DVA-access *and* explicitly reported to make use of such functionalities, (2) households that owned such devices *without* making use of built-in DVA-functionalities, and (3) households that did not even own such devices according to their survey responses. In fact, most households in the sample (47.8%) belonged to the second category, while the first category of owners *and* users only constituted the second largest group (28.4%), followed by the third category (23.8%) of non-owners *and* non-users (I in **Appendix XI**). Therefore, when it came to the characterisation of actual DVA-presence within childhood environments, this differentiation had to be taken into account. Interestingly enough, within each category, most households displayed average attitudes across all three attitudinal survey instruments, therefore implying that differences in DVA-ownership and usage did not necessarily translate into extreme attitudinal characteristics of the respective respondent. For instance, among DVA-owners *and* users (i.e. first category), 68.4% of households displayed average attitudes towards digital technologies and media in general, and 52.6% displayed an

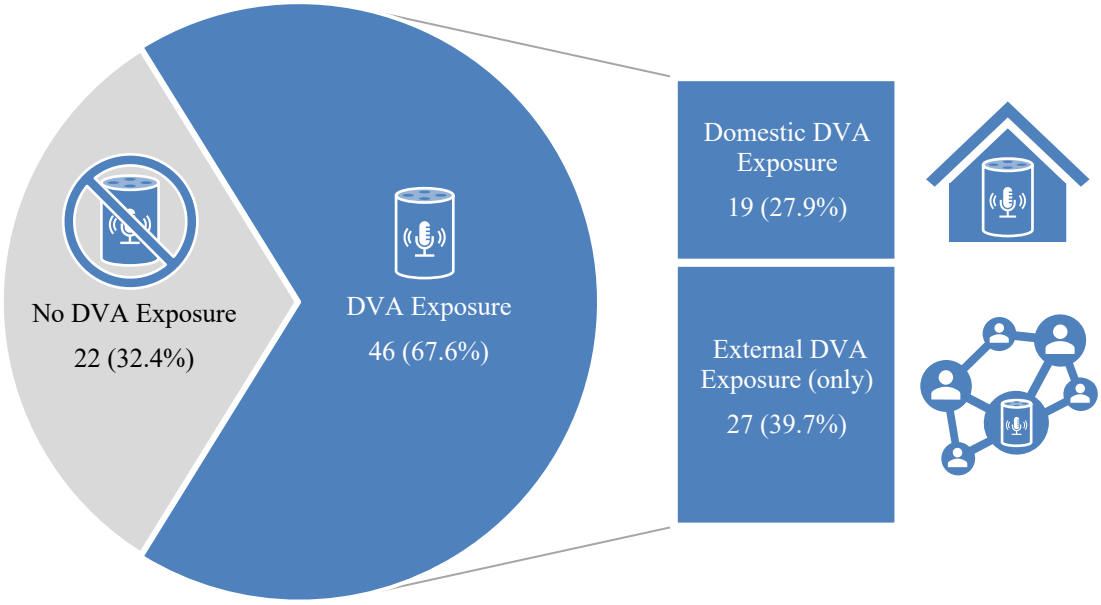
average attitude towards MI in particular (II in **Appendix XI**).[19] In addition, the different categories of DVA-ownership and usage were also spread across respondents' educational backgrounds: for instance, an equal number of DVA-owners and users in the sample had either completed a university degree (36.8%) or an apprenticeship (36.8%). This implies that DVA-ownership and usage did not cluster around a single level within the educational hierarchy.

The second main result of the survey data analysis was the important differentiation between children's *domestic* and *external* DVA-exposure: while domestic DVA-exposure was assumed if parental survey respondents reported ownership *and* usage of DVAs at home (i.e. first category following the categorisation used above), the survey also asked parental respondents of the other two categories whether their child had DVA-exposure in other contexts (e.g. while staying with relatives or friends). According to the data, a large proportion of parental respondents (39.7%) did indeed confirm that despite the absence of DVA-usage at home their child could actively engage with DVAs in other contexts or at least passively observe such engagements.

If both of these insights are taken into account, the presence of DVAs in childhood environments can be characterised in terms of *actual DVA-exposure*, which includes both, children who either had prior experience with DVAs at home (i.e. domestic) or within their closer circle of family and friends (i.e. external). In the context of the given study (**Figure 7**), this resulted in the overall finding that the vast majority of children (67.6%) already had been exposed to DVAs within their childhood environments, either domestically (27.9%) or externally (39.7%).

[19] A similar pattern became manifest in the mean values of attitudinal survey instruments: mean values of all three attitudinal scores differed by less than ½ SD across the total sample and all categories of DVA-ownership and usage (**Appendix XII**). Although there were some mean differences that would be worth investigating under the conditions of inferential statistics, especially regarding attitudes towards MI, this could not be done in the given context due to the non-random nature of the sample (**2.4.2**).

Figure 7. Presence of DVAs within childhood environments in the total sample

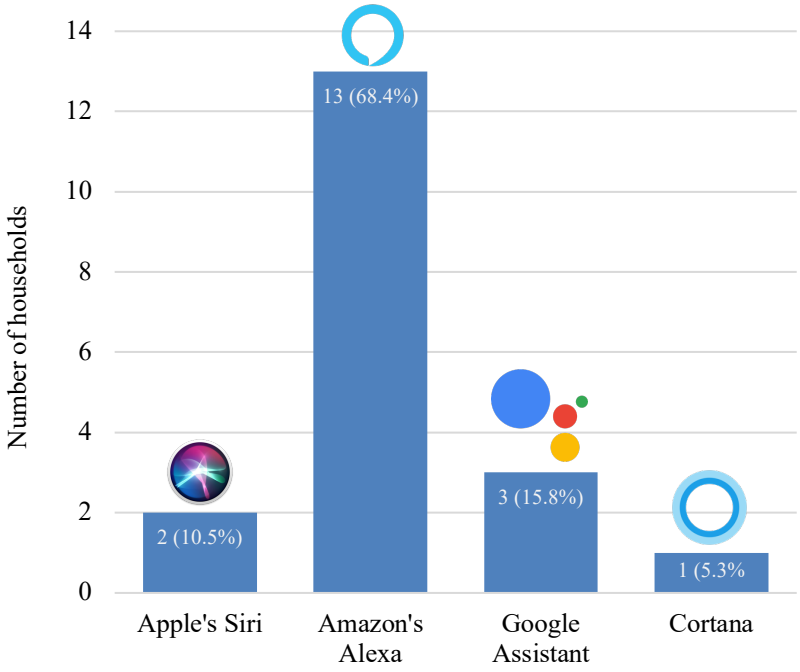


Notes. Figure visualises the ratio of children with (n=46) and without (n=22) DVA-exposure as well as the sub-ratio of domestic and external DVA exposure. 'Domestic DVA-Exposure' (n=19) means that according to the parental survey respondent the household owned *and* used at least one DVA. 'External DVA Exposure (only)' (n=27) means that according to the parental survey respondent the household did not use any DVAs, but the child did have the opportunity to engage with DVAs in the closer circle of family and friends.

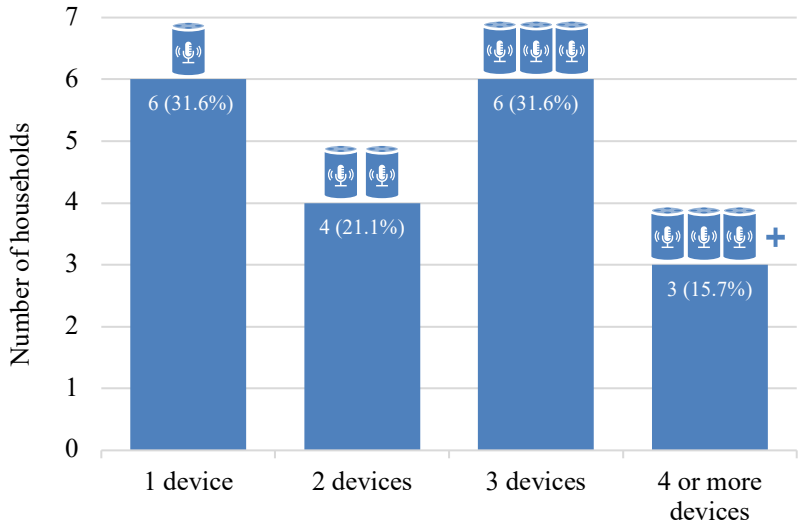
When it came to domestic exposure in particular, the survey data provided several additional insights into the presence of DVAs within household environments: firstly, Amazon's Alexa remained by far the most popular DVA, since it was used by 68.4% of households in the group of DVA-users (I in **Figure 8**). Secondly, DVAs were also quite spread within domestic childhood environments (II and III in **Figure 8**): 68.4% of parental respondents reported that DVAs were used on more than one device (21.1%: 2 devices, 31.6%: 3 devices, 15.7%: 4 or more devices), while popular locations included living rooms (68.4%), children's rooms (42.1%), dining rooms (21.1%) or kitchens (26.3%). Thirdly, the duration of domestic DVA-exposure further indicated that we did observe a very recent phenomenon, because 73.7% of households reported that they had started using DVAs less than 2 years ago (IV in **Figure 8**). Lastly, according to reported usage patterns (IV in **Figure 8**), listening to music, audiobooks or other types of audio entertainment were the most popular DVA-functionalities (89.6% used them at least 'occasionally'), followed by various features for daily practical assistance (74.0% used them at least occasionally) and interactive entertainment, such as games (63.2% used it at least 'occasionally'). In contrast, controlling smart home devices remained the least common DVA-feature (73.8% used it 'rarely' or 'never').

Figure 8. Domestic DVA-environments and usage patterns in the group of DVA-users

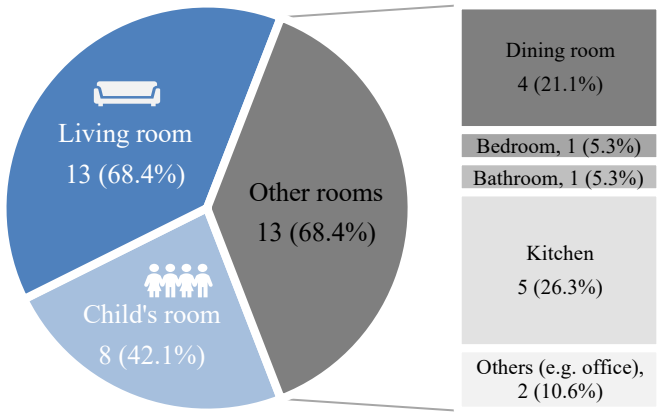
(I)



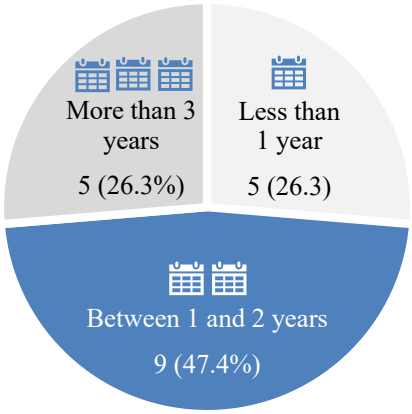
(II)



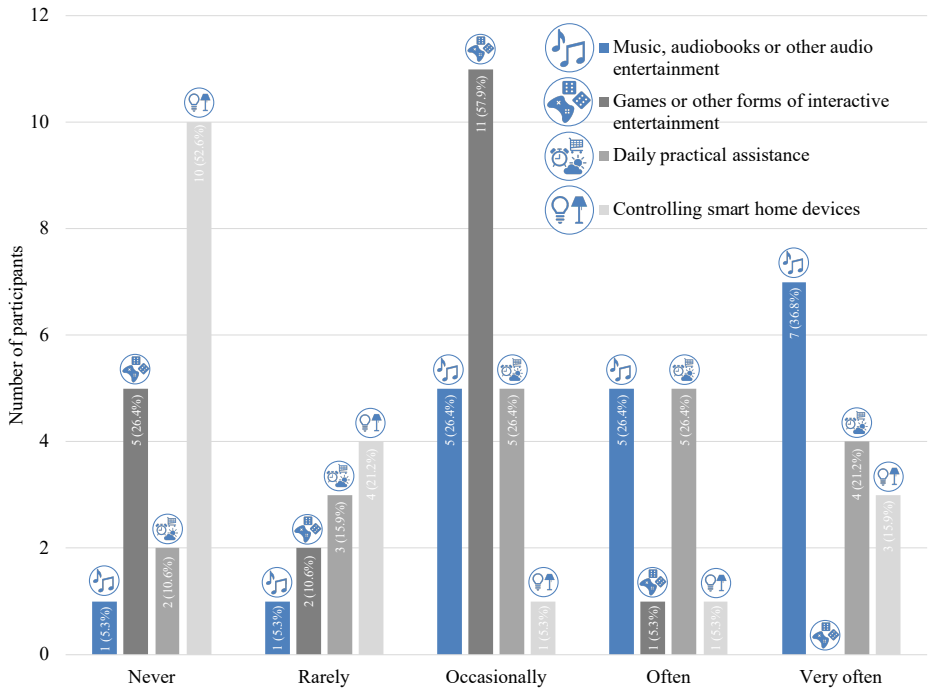
(III)



(IV)



(V)

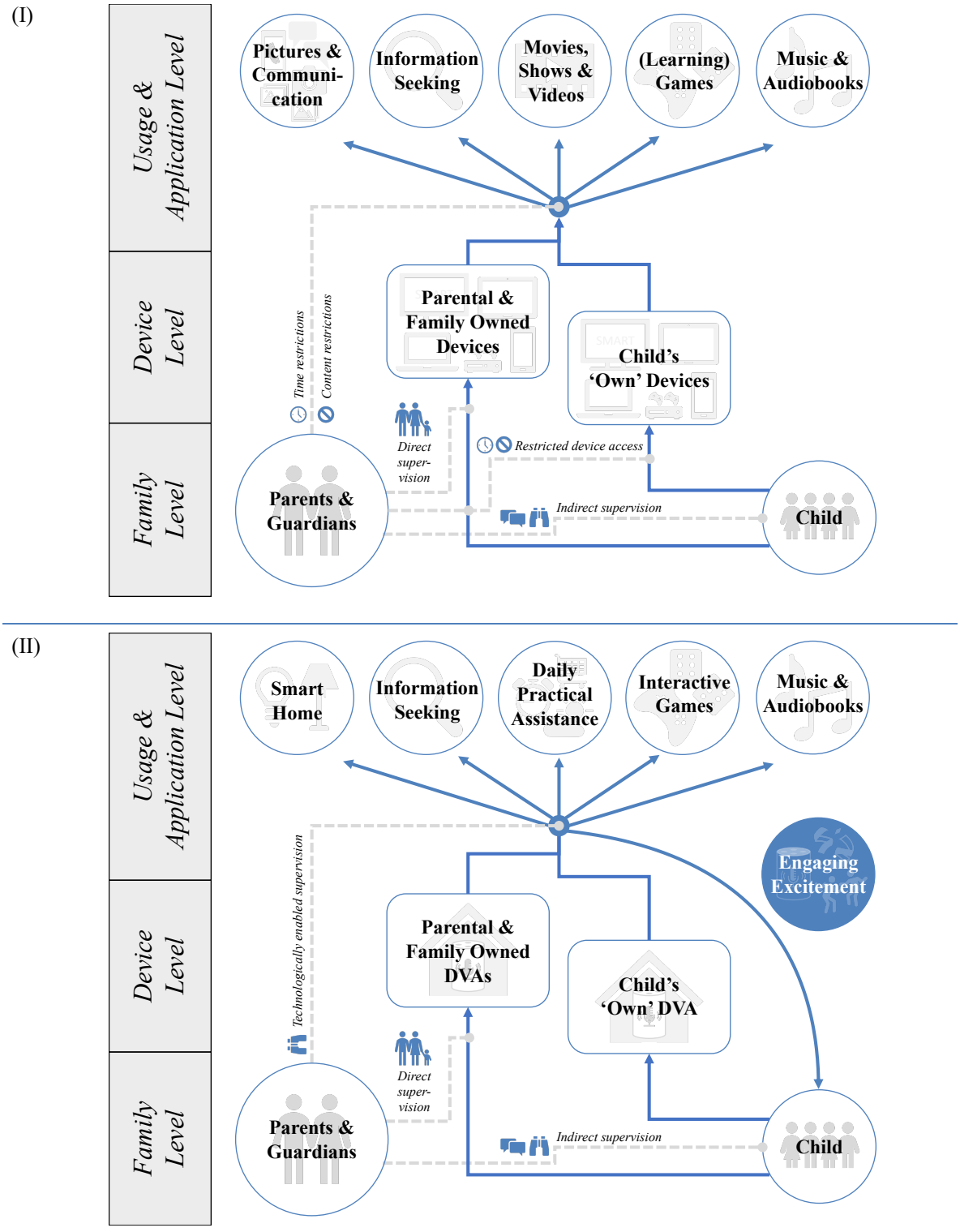


Notes. Figures (I) – (V) visualise domestic DVA-environments and usage patterns in the group of households that own *and* use at least one DVA (*n*=19). Figure (I) shows the presence of different DVA-brands in the sample. Figure (II) visualises the number of DVA-devices per household. Figure (III) visualises how long households had already used the respective DVA. Since none of the households in the sample reported a usage duration between 2 and 3 years, this entry option is omitted from the visualisation. Figure (IV) visualises where households use DVAs at home (sum of percentage values does not equal 100% due to multiple choice entry). Figure (V) visualises how often households use DVAs for different purposes (sum of percentage values equals 100% within each level of usage intensity). Households that used multiple DVAs were asked to answer questions with respect to the DVA that was used the most within the family.

(2) DVAs' embeddedness in the technological landscape of childhood environments

If we ignore DVAs for a moment, the thematic analysis of open entry survey responses (code book in **Appendix XIII**) suggested that children's general engagement with digital technologies and media was an omnipresent yet miscellaneous aspect of domestic childhood environments (I in **Figure 9**).

Figure 9. General and DVA-specific engagement with digital technologies and media



Notes. Figure (I) visualises how parental respondents in the sample (n=68) described their child's general engagement with digital technologies and media at home. Visualisation is based on aggregated results of a thematic analysis (part I in **Appendix XIII** for code book) that was applied to open entry item 25 (**Appendix IV**). Figure (II) visualises how parental respondents who owned *and* used at least one DVA (n=19), or who reported that their child did have the opportunity to engage with DVAs in the closer circle of family and friends (n=27), described their child's engagement with DVAs. Visualisation is based on aggregated results of a thematic analysis (part II.A and II.B **Appendix XIII** for code book) that was applied to open entry item 27 and 28 (**Appendix IV**).

At the usage and application level, primary school children in the sample had access to various digital functionalities and services, including the streaming of media content (e.g. music, audiobooks, movies, shows, videos), gaming (e.g. video games or learning games), communication (e.g. exchanging text messages or media with family members or friends via messenger services), or general information seeking (e.g. using search engines to retrieve information from the internet). At the device level, many of these functionalities and services were accessed through the same device, especially mobile phones and tablets. But often households also offered several devices with different access options to digital functionalities and services; for instance, mobile phones, tablets, or smart TVs served as popular access points for the streaming of media content or information seeking, while video game consoles or laptops were often used in addition for gaming or educational purposes. At the same time, various households reported to have 'child-owned' devices (e.g. smartphones, tablets, or video game consoles), meaning that the child did not have to share the respective device with other family members. Lastly, at the family level, parental supervision was exercised through different means, including *direct* (e.g. personally observing child's engagement with digital technologies and media) and *indirect* means of supervision (e.g. having general awareness of children's engagement with digital technologies and media). But many parents also reported to use *technological means of supervision*, either based on content and time restrictions at the usage and application level (e.g. browser restrictions, 'YouTube Kids', child-friendly search engines, screen time reports), or based on controlled device access (e.g. devices without internet access, video game consoles with age-appropriate games, time restrictions for device usage).

In light of this engagement framework, the question became how things changed when DVAs entered the technological landscape of childhood environments (II in **Figure 9**). On the one hand, the survey suggested that DVAs integrated consistently into the generic engagement framework described above: at the usage and application level, they allowed children to access very similar digital functionalities and services, such as the streaming of media content (e.g. music, audiobooks), gaming (e.g. quiz games), or general information seeking (e.g. asking knowledge-based questions). In line with the usage patterns described in the previous section, children also used Alexa et al. for popular DVA-functionalities and services, such as daily practical assistance (e.g. setting alarms, weather forecasts, setting ambient sounds before going to bed). Interestingly enough, the consistency continued at the device level, where DVAs were either 'owned' by the child or shared within the family. In the group of family-shared DVAs, parental respondents often described engagements as social

happenings with multiple family members usually being present. Hence, at the family level, parental DVA-supervision was either exercised through *direct* (e.g. personally observing the child's engagement with DVA) or *indirect* means of supervision (e.g. having general awareness of the children's engagement with the DVAs). Lastly, some parents also reported to use technological means of supervision based on DVAs' voice protocols or usage histories that could be accessed through parental mobile devices.[20]

Therefore, from this perspective, DVAs simply allow children – just like any other digital device – to access different functionalities and services that would often be accessible through other means as well. However, some of the themes in the data suggested that children's engagement with DVAs should not be reduced to simple access points for digital functionalities and services: many parental respondents also reported high levels of *engaging excitement* when their child was exposed to DVAs, either domestically or within the closer circle of family and friends. Typical scenarios included: asking funny questions and laughing about DVAs' strange answers (e.g. "Alexa, how are you?"), using funny features (e.g. Alexa's so-called 'Fart Generator'), or similar probing schemes. Although according to some survey respondents this engaging excitement – despite being very strong in early phases when children had been newly exposed to DVAs – seeped away over time, such observations could be systematically different compared to other usage and application scenarios, because in these observed contexts children's behaviour might be entirely driven by the excitement to engage with a responsive entity via a humanoid voice and without accessing particular functionalities and services (II in **Figure 7**).[21]

3.1.2 Mixed Methods Focus Group Results

The rich and unstructured focus group data, as they were embodied by transcripts and audio-visual recordings, allowed for a further in-depth analysis of children's engagement patterns with DVAs as well as their potential embeddedness in ontological belief systems (code book in **Appendix XV**).

[20] Although Amazon recently endowed Alexa with a 'parental dashboard' that allows for various technological means of supervision (e.g. bedtime periods, daily time limits, or content filters; cf. Amazon, 2019), none of the parental survey respondents in the sample explicitly reported to make use of this specific functionality.

[21] On a side note, one should keep in mind that the majority of households in the sample (72.1%) still reported that they did not use DVAs at home. Survey results on reasons for refraining from using DVAs at home are summarised in **Appendix XIV**.

(1) Characterising children's general ontological belief system

On an aggregate level, the focus group data revealed rich insights into children's perception of the *human* category within their ontological belief system: children described humans as *living* beings and referred to various aspects of what it means to *live* or *be alive*, such as breath, physical motion, emotions, or sensations. While the discussion remained very body-centred in the beginning, children were also able to consider various additional dimensions of human nature after being probed by the SI (e.g. human individuality in terms of personalities, preferences, faces or voices). Older children even added that humans are mammals, that our evolutionary ancestors were apes, and that we can only survive as part of nature's ecosystem due to our reliance on food and oxygen. In addition, within the overarching domain of living entities, children were also able to point out other ontological categories, such as plants, micro-organisms and, most importantly, animals. When children were asked to compare humans with animals, they also collated several dimensions of similarity as well as dissimilarity, including the uniqueness of *human language*, which was mainly attributed to the use of signs and translatability. For instance, Jordan (M/10) argued:

> *The special thing about human language is that all humans can understand it.* [other children start shaking their heads] *Yes, I mean there are different languages, but among animals, languages are separated by species. Apes have their own language. Cats have their own language. And dogs. But humans can also speak different languages – if they are smart.*[22]

These initial hints of children's perception of HI were complemented based on further discussions that the SI probed by presenting the life-size model of the human brain (**2.3.1**). Children mainly characterised the brain as the human body's *control unit* (e.g. regulating bodily functions, processing sensations, triggering motions) and the *origin of thought* (i.e. humans need the brain to think). Furthermore, children across all age groups contributed to the discussion by pointing out different dimensions or applications of HI, such as *memory* and *knowledge* (e.g. memorising spellings, knowing information, knowing how to behave) or *mathematical thinking* (e.g. solving math problems, mental arithmetic, knowing how to write numbers). When it came to the characterisation of 'intelligent' or 'smart' people, the discussion not only re-emphasised the importance of knowledge (e.g. memorising lots of information and facts) but also *task orientation* (e.g. solving pre-defined tasks or knowing correct answers).

[22] Children's names are always presented with their respective age and biological gender in brackets. For instance, Jordan (M/10) indicates that at the time of the data collection Jordan was a 10-year-old male according to the parental survey respondent.

To draw an analytical interim conclusion, based on the findings outlined above we can state that the ontological framework introduced earlier (**Figure 1** in **1.3.3**) was reflected in the focus group data on an aggregate level, meaning that the ontological belief systems of children in the sample seemed to follow the framework's general categorical pattern – at least within the domain of living entities. But how did this translate into their ontological perception of machines and MI?

Initially, children mainly described machines (e.g. computers or robots) as *memory systems*, because they could store or access information in reliable ways without forgetting them. In particular, the choice of words that children used to describe machines also confirmed their firm *artefactual perception*, such as 'turning them on or off', 'hard drives', 'hard disks', 'computer chips' or, most importantly, 'programming'. Indeed, the perception of machines as *programmable* entities was another very prevalent theme, meaning that children explicitly stated how we as humans provide the essential inputs to computers, robots or even DVAs.

This perception also gave rise to a very interesting discussion regarding the nature of the ontological human-machine relationship and the fundamental question whether MI can exceed HI. Although some children did mention *analogies* between the human brain and the computer (e.g. "Our brain is something like our internal storage disk."), the majority strongly defended the basic belief that HI must always be greater or equal to (≥) MI (**Table 8**). A prevailing reasoning among this particular group of children was that humans always have to 'put' their own intelligence into the machines they build or program, and therefore the intelligence of the latter entities could never exceed the intelligence of the former ones. Corresponding survey data revealed that parental respondents displayed either average or negative attitudes towards digital technologies and media in general (I in **Table 8**), MI (II in **Table 8**), or children's engagement with digital technologies and media (III in **Table 8**). Hence, children's firm belief about the ontological 'HI ≥ MI' hierarchy could indeed be driven by the absence of progressive parental attitudes towards technology. Although a few children did oppose such claims and indicated that MI would not necessarily be restricted by the particular HI who built or programmed it, therefore implying a flexible perception of the 'HI </>/= MI' hierarchy, none of them could elaborate on the exact reasonings, especially due to a lack of focus. Therefore, the data did not allow for a direct comparison of such opposing views.

Table 8. Firm beliefs about ontological hierarchy: 'HI ≥ MI'

Child	Gender/Age	DVA Exposure	Parental Attitudes (I)	(II)	(III)	Quotes
Jordan	M/10	No (-)	0	0	-1	*Computers cannot think like a human brain. And they always have to be programmed, unlike the human brain.* […] *That's why they can never do something on their own. Everything has to be programmed.*
Eva	F/10	Yes (domestic)	0	0	0	*A computer is a thing, a machine invented by humans. And it is endowed with intelligence from the human brain.* […] *And unlike computers, humans can always learn new things on their own, which then can be stored in computers. That's how they* [computers] *learn.*
Emily	F/10	Yes (external)	0	0	0	*Well, where else should a computer receive information from? Somebody has to put things in after all. They don't simply fall from the sky. That doesn't work.*
Jonas	M/9	Yes (domestic)	0	0	0	Jonas: *A computer can only be as intelligent as the human who invented it.* SI: *Okay, so who else agrees with this statement?* [to the rest of the group; but only a few children raise their hand] Jonas: *What, seriously? The rest of you doesn't think so?* [to the rest of the group; very shocked]
Kate	F/9	Yes (external)	0	-1	-1	*Technology can never be as good as the human brain.*
James	M/8	Yes (domestic)	-1	-1	0	*A robot for instance has to be programmed by a human. So, the human has to tell the robot what it is, because the robot doesn't know itself. But it knows, once we have put our own intelligence in it.*

Notes. Table summarises background information and exemplary quotes from children who argued that there is a strict ontological hierarchy between HI and MI, meaning that the latter one could never exceed the former one. Background information are based on survey data: age (in years), DVA-exposure at home (domestic) or within the closer circle of family in friends (external), categories of parental attitudes towards (I) digital technologies and media in general, (II) towards MI, and (III) towards children's engagement with digital technologies and media. Attitudinal categories range from -1 to 1 and were generated based on differences between individual attitudinal scores and the overall sample mean of the attitudinal score (i.e. low/high attitude [-1/1]: individual attitudinal score is below/above 1 SD of sample mean; average attitude [0]: individual attitudinal score is within 1 SD of sample mean).

(2) Characterising behavioural and verbal engagement patterns with DVAs

Children's DVA-engagement during focus groups revealed various consistencies with the survey results outlined earlier (**3.1.1**): at the usage and application level, children with previous DVA-exposure mainly accessed those digital functionalities and services that also prevailed among parental respondents, such as the streaming of media content (e.g. music, audiobooks), practical assistance (e.g. weather forecasts, time), information seeking (e.g. asking knowledge or trivia questions), or other interactive features (e.g. asking for jokes) – and especially some of the older boys (i.e. 9 – 10 years) used DVAs to ask for explicit media contents, such as obscene rap songs.

In addition, focus group data also showed strong support for the notion of e*ngaging excitement* discussed earlier (**3.1.1**). Within this theme, the most prominent pattern of engaging excitement was to evoke obscene responses from DVAs (e.g. several boys across all age-groups asked DVAs to make farting or burping noises, and one 10-year-old girl asked whether it could find her a boyfriend on a dating platform). Additional patterns of engaging excitement included various notions of playfulness. For instance, the following dialogue, which yielded to an amusing misunderstanding, occurred between Freya (F/9), Alice (F/6), Elisa (F/6), and Alexa:

Freya (F/9):	*Alexa, how do you like Mr.* [surname of the SI]?
Alexa:	*I do not have an answer, unfortunately*!
Alice (F/6):	*Mr.* [surname of the SI] *is a human!*
	[pause]
	Alexa, Mr. [surname of the SI] *is a human!*
Alexa:	*I am not sure about that, unfortunately!*
	[group bursts into laughter]

In another situation, the same group of girls engaged with the Google Assistant in several frisky dialogues, such as the following one:

Alice (F/6):	*Hey Google, shall I kiss you?* [leaning over closely and making a kiss gesture towards the Google Assistant]
Elisa (F/6):	*Should I tell her, that you just kissed her?* [to Alice]
Alice (F/6):	*Noo! Hey Alexa … ehh. Hey Google, Elisa just kissed you!*
Elisa (F/6):	*Noo! Hey Google, Alice just kissed you!*
Google:	*Sorry*, *I did not understand that.*
	[…]
Alice (F/6):	*Hey Google, you look really nice!*
Google:	*I am more than happy to return this compliment!*
Alice (F/6):	*Eh, you don't even know how I look like do you?*
	[pause]
	Hey Google, do you even know how I look like?
Google:	*Sorry, I can't help you with that yet!*
Alice (F/6):	*Eh, I think I already know how I look like!* [to the rest of the group]

This and other misunderstandings or playful dialogues typically caused peals of laughter among children, sometimes culminating in enormous noise levels so that the transcription of audio-visual recordings was almost impossible. Although some parental respondents in the survey reported that the intensity of engaging excitement might seep away over time, the focus group observations were quite consistent across various degrees of prior DVA-exposure. For instance, Alice (F/6) and Elisa (F/6) only had limited external or even none prior DVA-exposure according to survey data *as well as* their own verbal self-reports, but at the same time, their levels of engaging excitement appeared to be very similar to other children who had more than 3 years of domestic DVA-exposure according to survey data.

Furthermore, the dynamics of children's observed engagement patterns with Alexa and the Google Assistant showed great *volatility*, because children either interfered in each other's voice commands by 'stealing' wake words (e.g. person 1: "Alexa, …" ⇆ person 2: "… turn the radio on!") or rapidly introduced new voice commands (e.g. music songs never played for more than 10 seconds during the focus groups).

Lastly, the tone of children's engagement with DVAs varied tremendously: on one extreme end, the dialogues quoted above – and several others not quoted here – between Alice (F/6), Elisa (F/6), and the DVAs represented the most *affectionate* and *friendly* tones that occurred in the data. In addition, their engagement was not only characterised by such extreme patterns of affection, but they also used more subtle cues related to courtesy and politeness (e.g. using 'please' or 'thank you' while engaging with the DVAs). In contrast, their older group mate Freya displayed the other extreme end of the spectrum, namely *unfriendly* or even *aggressive* behaviours (e.g. using swear words, insults and very dominant verbal or body language towards DVAs). For instance:

Freya (F/9):	*Alexa, what's the time?* [looking at her own wristwatch simultaneously]
Alexa:	*It's 13:14 …*
Freya (F/9):	*… 15, you stupid thing! This is a radio clock!* [points at her wristwatch] *You are a stupid …*
Alice (F/6):	*… noo!* [trying to hold her hand over Freya's mouth]
	Alexa, you are really great! [to Alexa]
Elisa (F/6):	*Yes Alexa, you are really great!* [to Alexa]

Although only a few other children across age groups displayed such extreme behaviours towards DVAs, on an aggregate level the general tone of children's engagement with DVAs

did remain rather *plain*, *dominant,* and *unfriendly* in the sense that they differed significantly from the tone of children's observable intra-human engagements (e.g. among each other or with the SI during focus group discussions). For instance, many children used very dominant verbal expressions to pause DVAs (e.g. "Hey Google, be quiet!", "Alexa, shut up!", "Alexa, shut your face!"). Therefore, it still remains an interesting yet limited observation that Alice (F/6) and Elisa (F/6) were the only two exceptions who stood out in terms of affability and affection, because both of them maintained an unambiguous friendly stance towards Alexa and the Google Assistant. Interestingly enough, as mentioned above, both of them were one of the few children in the focus group sample who only had limited external or even none prior DVA-exposure.

(3) Characterising the probing schemes of DVAs' intelligence and identities

The focus group data allowed for the identification of several themes that children used in order to probe DVAs' perceived intelligence and identities. Firstly, the probing of DVAs' *knowledge* stood out in terms of comprehensiveness, because this theme itself comprised 6 characteristics. Among them, asking for nominal facts (i.e. questions which can be answered with one word or sentence, such as "Alexa, do polar bears live at the South Pole?") and quantities (i.e. questions that ask for a number, such as "Alexa, how many people live in Asia?") were the most prevalent probing schemes, which occurred across all age groups and regardless of prior DVA-exposure. A third characteristic within the overarching knowledge theme referred to the meaning of words and names (e.g. Jordan: "Alexa, what does Jordan mean?" or "Alexa, what is an argument?"). In contrast, an equally prevalent characteristic was to ask open questions (i.e. questions which cannot be answered with one word or sentence); for instance, Alice (F/6) and Elisa (F/6) asked DVAs to tell them something about America, Africa, the North Pole, the primeval forest or, most interestingly, how a human being looks like (see below). Alice (F/6) also asked those questions which formed the fifth knowledge-related characteristic: interpersonal knowledge. For instance, she tested whether Alexa knew who her favourite teacher in school was, or whether the Google Assistant knew what her father's name. The fact that Alice (F/6) was the main driver behind the last characteristics (i.e. open questions and interpersonal knowledge) is another interesting observation, because as pointed out earlier she was one of the few focus group participants without any prior DVA-exposure according to survey data as well as her own verbal self-reports. The last knowledge-related characteristic comprised voice commands that probed a variety of basic skills, including rather generic skills (e.g. "Hey Google, say a word!", Leon:

"Hey Google, say *Leon*!", Megan: "Hey Google, how do you spell *Megan*?"), mathematical skills (e.g. "Hey Google, count to 10!", "Hey Google, what is 2000 + 2000?"), or foreign language skills (e.g. "Hey Google, can you say a word in Spanish?", "Alexa, count to 20, but in English!").

Children also explored DVAs' *identities* by asking a variety of questions related to personality traits or background information. While many of these questions referred to rather human, or, for that matter, humanoid characteristics (e.g. "Alexa, how old are you?", "Hey Google, do you have a girlfriend?", "Hey Google, what is your name?", "Alexa, are you married?", "Alexa, what gender are you?"), some children also asked questions of artefactual nature (e.g. "Alexa, tell me something about how you were built?", "Alexa, who built you?", "Alexa, do you have eyes?"). Interestingly enough, the latter questions were, again, only asked by Alice (F/6) and Elisa (F/6), therefore suggesting that even very young children with limited or no DVA-exposure seemed to be aware of DVAs' artefactual, that is, humanly made ontological nature.

Besides exploring DVAs' knowledge and personality traits, children also tested to what extent the machines could answer questions related to reflexivity (i.e. questions that usually require answers with some introspective content). On the one hand, some of these questions were of rather ordinary nature (e.g. "Hey Google, how are you doing?", "Alexa, what are you doing?", "Alexa, are we friends?"). On the other hand, in the special context of the HCTTG and the instruction to figure out whether DVAs might be controlled by a human, some children also asked very specific questions in order to shed light on DVAs' underlying ontological nature (e.g. "Hey Google, who are you, really?", "Alexa, where are we right now?", "Alexa, are you controlled by a human?" or "Hey Google, are you a human?"). Obviously, Alexa as well as the Google Assistant often had a ready reply to such questions, but the last section shows how children interpreted DVAs' response behaviours in very unique ways.

(4) Perceptions of DVAs in terms of humanoid and non-humanoid characteristics

Focus group discussions centred around the HCTTG were supposed to elicit what perceived DVA-characteristics children interpreted as being genuinely *humanoid* (i.e. attributable to alleged human controller) or *non-humanoid* (i.e. attributable to DVA itself).

For *non-humanoid* characteristics, three different categories of themes were identified in the data: firstly, regarding the dynamics of DVAs' response behaviours, children perceived the

devices' *promptness* and *permanent responsiveness* as very non-humanoid characteristics, meaning that they did not assume a human controller in the background if DVAs displayed such response behaviours; this mostly occurred when children probed DVAs' knowledge by asking for nominal facts or quantities. For instance, like most children Lukas (M/9) thought the Google Assistant would be controlled by a human and argued:

> *Alexa can filter information from the internet so fast, it's just impossible for a human to do so at the same speed. And when we asked Alexa, she just answered right away, and much faster compared to Google. That's why we think Google was controlled, because he* [the alleged human controller] *probably had to look things up first.*

This argument was reflected in one of the DVA-engagements Lukas (M/9) had experienced earlier during the focus groups:

Lukas (M/9):	*Alexa, how many people live in Asia?*
Alexa:	*Asia has 3.879.000.000 inhabitants.*
Lukas (M/9):	*Okay, seriously, nobody could answer this so quickly! This is definitely the real Alexa! Do we agree on that?* [to the rest of the group]

Similarly, Jordan (M/10) and Kate (F/9) argued:

Jordan (M/10):	*As a human, you just can't be so quick, going from A to B, answering questions and then playing music. I mean, it took like one second and then the music was there!* [referring to Alexa]
Kate (F/9):	*Yes, it just took Google way too long to come up with answers. And Alexa, she could answer like a bullet from a gun.* [typical idiomatic expression in the German language] *And Google often just went off and did not answer at all!*

Kate's (F/9) statement was also related to another aspect of DVAs' response dynamics that some children perceived as a rather non-humanoid characteristic, namely DVAs' *non-responsiveness*. However, the interesting detail is that only two children, namely Alice (F/6) and Elisa (F/6), perceived this non-responsiveness as a non-humanoid characteristic, while the majority – including Kate – interpreted the very same observation as a strong humanoid sign (see below).

The second category of perceived non-humanoid characteristics referred to children's experienced conversational quality with DVAs: whenever the machines displayed a *lack of common sense* or *struggled to understand* children, this was usually interpreted as 'true' DVA-behaviour. For instance:

Emily (F/10):	*Alexa, are there polar bears at the South Pole?*

Alexa:	*I am not sure about that, unfortunately!*
	[…]
Lukas (M/9):	*Guys, first of all, a human would definitely understand this question!*
Emily (F/10):	*Yes, and also know the answer!*
Lukas (M/9):	*No, I mean 'understand'! He would 'understand' the question! And maybe Alexa just didn't understand us properly and that's why she couldn't answer it.*

Similarly, Alice (F/6) and Elisa (F/6) surprised the SI with a remarkable argument regarding the Google Assistant's lack of common sense:

Elisa (F/6):	*We asked* [the Google Assistant] *how a human looks like!*
Alice (F/6):	*And she didn't know!*
Elisa (F/6):	*Yes, she said something like 'Unfortunately, I do not have a screen to show you'.*
SI:	*Well, but doesn't that make sense? I mean there is no screen, right?*
Elisa (F/6):	*Yes, but a human knows how a human looks like and could describe it at least!*

Lastly, the third category that children used in order to identify non-humanoid characteristics was based on the concrete nature of DVA-responses. On the one hand, children with prior exposure to DVAs were *familiar with the exact wordings* of standardised answers (e.g. "I do not have an answer, unfortunately!", "Shuffling your Music on Amazon Music!") and used this familiarity to conclude that they were conversing with the 'true' DVA instead of a human controller. On the other hand, children also interpreted the *accuracy* of DVA-responses as a very non-humanoid characteristic. For instance, during the group discussion, Lukas (M/9) referred back to the DVA-conversation quoted earlier and stated:

> *So, we asked Alexa 'How many people live in Asia?' and then she said … ehh … I already forgot the number.* [laughs] *But anyway, it was very convincing, because she even specified the hundreds number.*

In turn, children perceived *inaccurate* DVA-responses as a very *humanoid* characteristic, for instance if it reacted mistakenly:

Jordan (M/9):	*Alexa* [Google Assistant lights up], *ehh I mean …*
Jonas (M/9):	*It went on! See! It still went on! That's him* [alleged human controller]*, he made a mistake!*

Furthermore, children also referred to the dynamics of DVAs' response behaviours in order to identify humanoid characteristics: as indicated above, *non-responsiveness* was perceived as a very humanoid sign by most children:

Kate (F/9):	*Hey Google* [Google Assistant lights up]*, who are you, really?*
Stephanie (F/8):	*Seems like somebody is thinking right now what to answer.*
	[Google Assistant switches off]
Max (M/8):	*Look, it turned off!*
Stephanie (F/8):	*It's him* [alleged human controller]*!*

Another non-response was also interpreted in a similar way, but due to a different reason:

Emily (F/10):	*Hey Google* [Google Assistant lights up], *don't you think Jeff is an asshole?*
Jeff (M/10):	*You are one yourself!* [to Emily]
	[Google Assistant switches off]
Lukas (M/9):	*Hey, look at that! I think he* [alleged human controller] *turns it off whenever you say something like that!*

Hence, Lukas (M/9) perceived the fact that the Google Assistant switched off as a very humanoid move to ignore Emily's insulting question, further implying that for him such acts of disregarding reprehensible statements must come from an actual human being rather than a machine.

Additionally, and in correspondence with the findings on *promptness* reported above, most children perceived *delayed responses* as an indicator for human control. For instance, Emily (F/9) argued:

> *We wanted to listen to a song, and Alexa could immediately play it. And this one* [pointing at the Google Assistant] *took forever. I guess he* [alleged human controller] *tried to look up the song, but then couldn't find it after all, because it said something like 'Sorry, I am not sure whether I understood you correctly!'*

In contrast, as pointed out above, Alice (F/6) and Elisa (F/6) argued in the opposite direction by claiming that Alexa's *promptness* and *permanent responsiveness* constituted a very *humanoid* characteristic. This 'contradiction' remains a very interesting observation, because as mentioned before they are both one of the few children who only had limited external or even none prior DVA-exposure according to survey data as well as their own verbal self-reports. This could indicate that – cautiously spoken – DVA-exposure does change children's interpretive perception of certain behavioural patterns displayed by intelligent machines.

Lastly, regarding the concrete nature of DVA-responses, children mostly perceived their pre-programmed repertoire of *creative verbal outputs*, such as puns or very playfully formulated answers, as a rather humanoid characteristic:

Stephanie (F/8):	*Hey Google, how are you doing?*
Google:	*I feel positively charged! What do you want us to do?*
	[children start laughing]
Kate (F/8):	*Wait a minute. Please, be quiet guys. I've never heard Alexa or Siri saying 'charged'! That must be him* [alleged human controller]*!*

Other groups of children experienced similar situations, such as:

Jeff (M/10):	*Hey Google, do you have a girlfriend?*
Google:	*I do feel closely connected to the internet, but the most interesting conversations are the ones I have with you!*
	[children start laughing]
William (M/10):	*I bet that was him* [alleged human controller]*! A hundred percent. He just typed this answer!*

3.2 Discussion of Results

The final step remains to integrate the rich array of mixed methods findings reported above into one discussive strand that can provide answers to this inquiry's guiding research questions and contribute to the research gap identified earlier (**1.4**).

Firstly, according to the results, the nature of today's childhood environments *does* change: if domestic as well as external DVA-exposure are taken into account, one should suspect that the vast majority of today's children grows up amid Alexa et al., which means that they either actively engage with DVAs themselves, or they passively observe such engagements within their environment. Therefore, the implicit claim of this discussion's title can be confirmed on empirical grounds, because given that Alexa et al. were only introduced a couple of years ago, the observed presence of DVAs must constitute a 'true' change. Furthermore, and in line with recent German market research (e.g. Postbank, 2019; Quantcast, 2019), the results suggest that this change is *homogenous* in the sense that it is mainly driven by two voices, namely Alexa and the Google Assistant. At the same time, this change also seems to be *ubiquitous* in the sense that DVAs are not only spread *within* domestic environments (i.e. used on multiple devices) but also *across* different family environments (i.e. across educational backgrounds or attitudes towards digital technologies and media).

Secondly, within this context of *changing* childhood environments, the nature of children's engagement with DVAs remains a twofold phenomenon: on the one hand, children not only use DVAs to access similar functionalities and services compared to other types of digital devices, but they also experience similar means of parental supervision, which mainly depend on whether the DVA is 'owned' by the child or shared within the family. In the latter case, the results are also consistent with previous research that showed how DVAs can become well-integrated devices of daily family life by providing various functionalities and services through their unitary voice-user interfaces (cf. Porcheron et al., 2018). However, this viewpoint would barely suggest – in line with the wording from the introduction – that DVAs could constitute a profound change with historical momentum, because the presence of DVAs would be reduced to a novel type of interface within the existing landscape of digital technologies and media.

But on the other hand, results show how communicating with DVAs via humanoid voice-user interfaces seems to be a particular source of *engaging excitement* for children, further implying that in these scenarios the engagement is not a means to an end but rather an end in itself. This finding does not only emerge from survey data but also from focus group data: children display various behavioural and verbal engagement patterns with DVAs that convey strong notions of playfulness, therefore implying that they might perceive DVAs as actual playmates and engagement partners in their own right – notably, regardless of previous DVA-exposure. This finding also follows the general notion reported in the empirical literature on child-machine engagements (**1.1**), which states that children's engagement with intelligently behaving machines can yield to patterns similar to the nature of intra-human engagements. To a certain extent, this is also consistent with Druga et al. (2017), who not only report various patterns of engaging excitement (**Table 2** in **1.4**) but also rather friendly and affable tones that children use while engaging with DVAs.

In contrast, the given inquiry yields a mixed empirical picture regarding those tones of engagement, because results show both, friendliness or affability on the one hand and unfriendly or even aggressive tones on the other hand. Although the extremes on both ends might be exceptions, the general tone that prevails is still skewed towards rather plain, dominant and unfriendly voices that children use in order to communicate with DVAs. This suggests that the majority of children does not perceive DVAs as morally valuable or respectable entities comparable to human beings, which could also explain other empirical patterns in the data: if children feel less obliged to adhere to those moral norms that normally guide their intra-human engagements, it is neither surprising that engaging excitement often

implies children's appeal to evoke obscene behaviours and responses from DVAs, nor that they vigorously probe DVAs with rapidly changing voice commands.

But the empirically grounded claim that Alexa and the Google Assistant are not perceived as morally valuable or respectable entities comparable to human beings is not only relevant in terms of engagement patterns, but it can also be related to DVAs' perceived ontological status within the cognitive framework derived above (**Figure 2** in **1.3.4**): children hold firm ontological beliefs about humans' unique and also superior status within the domain of living entities, which translate into children's argument that humans as the creators of machines not only remain the masters over their created servants but also the limiting benchmark of machines' inherent capacities, meaning that MI could never exceed HI. Hence, children's empirically observable moral stance towards DVAs seems to be consistent with their ontological perception of the 'HI ≥ MI' hierarchy. In addition, as humanly made servants, DVAs' permanent responsiveness is often perceived as a non-humanoid characteristic, while children interpret non-responsiveness as a genuine human reaction of speechlessness at the same time. In other words: within children's ontological belief system, machines – as the servants of their masters – are created to respond at all times, while we as humans maintain our moral right to remain silent or to disregard others.

Another characteristic regarding DVAs' embeddedness in children's ontological belief system becomes manifest with respect to the concrete notion of intelligence: the results suggest that children's perception of HI seems to be driven by their own experience of what it means to be intelligent at an early stage in life, namely the daily struggles of learning how to read, write or speak languages, to acquire knowledge, to solve pre-defined tasks or to improve mathematical thinking. Children seem to apply this exact experience-based conceptualisation of intelligence when they probe DVAs through various questions related to knowledge or other skills that would prove the possession of intelligence when performed by themselves. In addition, the nature of observed probing schemes is also very consistent with the previous empirical findings in the literature (**Table 2** in **1.4**).

But despite holding firm beliefs about the ontological 'HI ≥ MI' hierarchy, most children are also convinced that at least within the knowledge-related domains of intelligence DVAs can outperform any human in terms promptness and accuracy, therefore suggesting that such characteristics seem to constitute MI's unique strengths in children's ontological belief systems. In turn, children perceive DVAs' limited conversational capacities as MI's unique weakness, because it proves the machines' lack of common sense and actual comprehension.

Furthermore, children often interpret creative responses – usually resulting from questions related to DVAs' identity or reflexivity – as an indicator for human interference, further implying that they do not associate this level of originality with DVAs' inherent capacities. Hence, according to both of these empirically grounded claims, children's ontological differentiation between HI and MI follows the belief that humans do remain superior in terms of creativity and actual verbal comprehension.

But do DVAs constitute *marginal* entities within children's emerging ontologies, that is, entities without a clear categorical belonging (**1.3.4**)? Firstly, one should note that children's empirically observable engagement with DVAs goes beyond the concept of anthropomorphism (**1.2**), because they do not engage in strong imaginative 'pretend-play' (i.e. 'as if the DVA was living'; cf. Airenti, 2015; Severson & Carlson, 2010). For instance, when DVAs pay compliments on appearance, even young children do not hesitate to express their doubts regarding the authenticity of the compliment rather than *pretending* that it was a genuine comment on their actual appearance. Secondly, children's perception of DVAs displays both, ontological firmness as well as openness: on the one hand, there seems to be a strong awareness of DVAs' artefactual nature as a humanly made and programmable entity. But on the other hand, results also indicate that children do not hesitate to interpret various aspects of DVAs' response behaviours as genuine acts of human interference. Moreover, the uniqueness of human language plays a dominant role in children's understanding of humanness in general and HI in particular, therefore suggesting that certain entities, such as DVAs, might be indeed able to testify their ontological marginality through the convincing mastery of human language.

However, does all this indicate – following the general claim of the NOCH (**1.3.4**) – that DVAs could constitute an ontological category in their own right? Although a conclusive answer is certainly beyond the scope of this exploratory inquiry, the signs of marginality discussed above point in an affirmative direction. Furthermore, children with limited or even none prior DVA-exposure (e.g. Alice and Elisa) can display very different – not say contradictory – behaviours and ontological perceptions compared to other children, therefore suggesting that DVA-exposure might indeed cause substantial changes in children's cognitive belief system – it *might*! But the questions of causality, and whether such changes only alter the perceived characteristics of *existing* categories or yield to a *new* one, remain subject to future research and more appropriate designs that allow for causal claims. Lastly, the results also raise further questions how children's perceptions of DVAs and respective engagement patterns develop with age, or across different genders and parental influences.

3.3 Reflexivity & Rigour (2): The *Interpretive* Dimension

Since some of the empirically grounded claims stated in the above discussion might appear overly determined and strong, it becomes all the more important to briefly assess their interpretive rigour according to Tashakkori & Teddlie's (2008) integrative framework.[23] Firstly, interpretive quality requires *consistency*, meaning that claims should emerge from empirical grounds in a stringent way and without departing from them (Tashakkori & Teddlie, 2008). For instance, the empirical claim that from children's perspective DVAs can constitute engagement partners in their own right (i.e. engaging excitement) gains credibility because it emerged from focus group as well as survey data and it was also found to be consistent with previous findings reported in the empirical literature. In contrast, the consistency of other claims can be questioned due to the absence of empirical counterparts in the data: for instance, if the association between children's negative moral behaviour towards DVAs and their ontological perception of the 'HI ≥ MI' hierarchy was truly consistent, it would imply that children who perceive the hierarchy differently should also display more egalitarian moral behaviours towards DVAs. Unfortunately, such corresponding counterparts were beyond the scope of the data.

Secondly, and most importantly, *integrative efficacy* assesses to what extent findings were rigorously integrated across methods into one coherent strand that should be more meaningful than the sum of its claims (Tashakkori & Teddlie, 2008). In the exploratory context of the given inquiry, triangulation between methods and data sources did not only allow for a comprehensive characterisation of the role that DVAs play in today's childhood environments, but triangulation on an individual level through data linkages also increased the depth of various insights that became manifest in the focus group data (e.g. **Table 8** in **3.1.2**). However, the overall inferential leverage of triangulation on an individual level was also limited by the absence of corresponding empirical counterparts in the data. For instance, only a few children shared opposing views regarding the ontological 'HI ≥ MI' hierarchy, therefore weakening the claim that such views could be determined by certain parental attitudes towards technology.

[23] Unfortunately, due to reasons of written space, an all-encompassing discussion of Tashakkori & Teddlie's (2008) framework on interpretive quality is beyond the scope of this section. Therefore, the focus remains on the two most important aspects in the given context: *interpretive consistency* as well as *integrative efficacy*.

Conclusion

In the aftermath of this long discussion, the reader's drained mind might be modestly satisfied with a concise conclusion: at its very core, and inspired by the principle of Kranzberg's Law, the present study intended to investigate to what extent DVAs might constitute a profound technological application with historical momentum. And although a conclusive answer was certainly beyond the scope of this exploratory inquiry, according to the findings – and despite some of their implicit methodological weaknesses – we should have reasons to believe that the growing presence of DVAs in today's childhood environments does translate into consequences that go beyond their immediate assistive purposes. In a nutshell, there are two essential arguments that can be raised based on the above discussion: firstly, even today children's engagement with DVAs is not always a means to an end but sometimes an end in itself, therefore suggesting that DVAs might be able to establish themselves as independent engagement partners during childhood and beyond. Secondly, as marginal entities, which embody humanoid as well as non-humanoid characteristics at the same time, DVAs are in a very powerful ontological position to leave their marks on children's cognitive development.

But to conclude this discussion with the greatest words of universal academic wisdom: *further research is needed*! Luckily, this inquiry is certainly not short of concrete starting points for the next steps, including further investigations of developmental trajectories (e.g. How do DVAs change children's ontological perceptions of HI as well as MI *within* individuals over time, or *between* individuals, for instance across genders and different parental influences?) or the effects of DVA-designs (e.g. How do differences in DVAs' voice-user interfaces affect children's ontological perceptions?). But in light of more than 100 million operating DVAs across the globe, the time is running …

References

Adler, P. A., & Adler, P. (1996). Parent-As-Researcher: The Politics of Researching in the Personal Life. *Qualitative Sociology*, *19*(1), 35–58.

Airenti, G. (2015). The Cognitive Bases of Anthropomorphism: From Relatedness to Empathy. *International Journal of Social Robotics*, *7*(1), 117–127. https://doi.org/10.1007/s12369-014-0263-x

AISB. (2019). AISBX 2019 / Loebner Prize. Retrieved June 26, 2019, from The Society for the Study of Artificial Intelligence and Simulation of Behaviour website: https://www.aisb.org.uk/events/loebner-prize

Amazon. (2019). Amazon Parent Dashboard. Retrieved January 24, 2019, from Amazon.com website: https://www.amazon.com/b?ie=UTF8&node=17395968011

Amazon Web Services. (2018). Alexa for Business. Retrieved November 15, 2018, from Alexa for Business Service website: https://aws.amazon.com/alexaforbusiness/

Anderman, C., Cheadle, A., Curry, S., Diehr, P., Shultz, L., & Wagner, E. (1995). Selection Bias Related to Parental Consent in School-based Survey Research. *Evaluation Review*, *19*(6), 663–674.

Anderson, C. (2019). Gradient Descent. In J. Brockman (Ed.), *Possible Minds: Twenty-Five Ways of Looking at AI* (pp. 143–150). New York, NY, US: Penguin Press.

Andreeva, V. A., Salanave, B., Castetbon, K., Deschamps, V., Vernay, M., Kesse-Guyot, E., & Hercberg, S. (2015). Comparison of the sociodemographic characteristics of the large NutriNet-Santé e-cohort with French Census data: the issue of volunteer bias revisited. *Journal of Epidemiology and Community Health*, Vol. 69. https://doi.org/10.1136/jech-2014-205263

Astington, J. W., & Hughes, C. (2013). Theory of Mind: Self-Reflection and Social Understanding. In P. D. Zelazo (Ed.), *The Oxford Handbook of Developmental Psychology, Vol. 2: Self and Other*. https://doi.org/10.1002/wcs.1232

Azevedo, R. (2015). Defining and Measuring Engagement and Learning in Science: Conceptual, Theoretical, Methodological, and Analytical Issues. *Educational Psychologist*, *50*(1), 84–94. https://doi.org/10.1080/00461520.2015.1004069

Bazeley, P. (2010). Computer-Assisted Integration of Mixed Methods Data Sources and Analyses. In A. Tashakkori & C. Teddlie (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research* (2nd ed., pp. 431–468). Thousand Oaks, CA, US: Sage Publications.

Bazeley, P., & Jackson, K. (2014). *Qualitative Data Analysis with NVivo* (2nd ed.). London, England, UK: Sage Publications.

Bernstein, D., & Crowley, K. (2008). Searching for Signs of Intelligent Life: An Investigation of Young Children's Beliefs About Robot Intelligence. *The Journal of The Learning Sciences*, *17*(2), 225–247. https://doi.org/10.1080/10508400801986116

Biesta, G. (2010). Pragmatism and the Philosophical Foundations of Mixed Methods Research. In A. Tashakkori & C. Teddlie (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research* (2nd ed., pp. 95–118). https://doi.org/10.4135/9781506335193

Bird, C. M. (2005). How I Stopped Dreading and Learned to Love Transcription. *Qualitative Inquiry*, *11*(2), 226–248. https://doi.org/10.1177/1077800404273413

Bjorklund, D. F. (2013). Cognitive Development: An Overview. In P. D. Zelazo (Ed.), *The Oxford Handbook of Developmental Psychology, Vol. 1: Body and Mind*. https://doi.org/10.1016/B978-0-12-397045-9.00059-8

Boyatzis, R. E. (1998). *Transforming qualitative information : thematic analysis and code development*. Thousand Oaks, Cal: Sage Publications.

Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Bricher, G. (1999). Children and qualitative research methods: A review of the literature related to interview and interpretive processes. *Nurse Research*, *6*(65–77).

British Educational Research Association. (2011). *Ethical Guidelines for Educational Research*.

Bryman, A. (2006). Integrating Quantitative and Qualitative Research: How is it Done? *Qualitative Research*, *6*(1), 97–113.

Bryman, A. (2016). The Nature of Quantitative Research. In *Social Research Methods* (5th ed., pp. 148–169). Oxford, England, UK: Oxford University Press.

Cameron, D., Fernando, S., Cowles-Naja, E., Perkins, A., Collins, E., Millings, A., … Prescott, T. (2017). Children's Age Influences Their Use of Biological and Mechanical Questions Towards a Humanoid. In Y. Gao, S. Fallah, Y. Jin, & C. Lekakou (Eds.), *TAROS'17 -18th Annual Conference Towards Autonomous Robotic Systems* (pp. 290–299). https://doi.org/10.1007/978-3-319-64107-2

Castells, M. (2004). *The Network Society: A Cross-cultural Perspective*. Cheltenham, England, UK: Edward Elgar.

Castro-González, A., Tan, X. Z., Carter, E., & Steinfeld, A. (2018). Social Haptic Interaction between Robots and Children with Disabilities. *ACM/IEEE International Conference on Human-Robot Interaction*, 81–82. https://doi.org/10.1145/3173386.3176984

Chow, M. (2010). Random Sampling. In N. Salkind (Ed.), *Encyclopedia of Research Design* (pp. 1210–1213). https://doi.org/10.4135/9781412961288

Coeckelbergh, M. (2009). Virtual moral agency, virtual moral responsibility: on the moral significance of the appearance, perception, and performance of artificial agents. *AI & SOCIETY*, *24*(2), 181–189. https://doi.org/10.1007/s00146-009-0208-3

Cohen, L., Manion, L., Morrison, K. (Keith R. B. ., Martin, J., Suri, H., & Martin, S. (2018). *Research Methods in Education* (8th ed.). London, England, UK: Routledge.

Cohen, M., Giangola, J., & Balogh, J. (2004). *Voice User Interface Design* (1st ed.). Boston, MA, US: Addison-Wesley Professional.

Cortina, J. M. (1993). What is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology*, *78*(1), 98.

Cowan, R. S. (1976). The "Industrial Revolution" in the Home: Household Technology and Social Change in the 20th Century. *Technology and Culture*, *17*(1), 1–23.

Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, & Mixed Methods Approaches* (5th ed.). Los Angeles, CA, UK: SAGE.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches* (4th ed.). Los Angeles, CA, US: SAGE Publications.

CUREC. (2019). CUREC Approved Procedures. Retrieved July 1, 2019, from Research Support website: https://researchsupport.admin.ox.ac.uk/governance/ethics/resources/ap#collapse397151

Dawar, N. (2018). Marketing in the Age of Alexa. *Harvard Business Review*, (May-June Issue), 80–86. https://doi.org/10.1108/sd.2010.05626gad.003

Day, B. M., Turner, G., & Drozdiak, N. (2019). Amazon Workers Are Listening to What You Tell Alexa. Retrieved April 16, 2019, from Bloomberg.com website: https://www.bloomberg.com/news/articles/2019-04-10/is-anyone-listening-to-you-on-alexa-a-global-team-reviews-audio

De Vaus, D. (2008). *Surveys in Social Research* (6th ed.). London, England, UK: Taylor and Francis.

DeForge, B. R. (2010). Research Design Principles. In N. J. Salkind (Ed.), *Encyclopedia of Research Design* (pp. 1252–1258). Thousand Oaks, CA, US: SAGE Publications.

Denzin, N. K. (1970). *The research act in sociology: A theoretical introduction to sociological methods*. London, England, UK: Butterworths.

Dreyfus, H. L. (1972). *What computers can't do: A critique of artificial reason*. New York, NY, US: Harper & Row.

Dreyfus, H. L. (1992). *What computers still can't do: A critique of artificial reason* (3rd ed.). London, England, UK: MIT Press.

Druga, S. (2018). *Growing Up With AI - Cognimates: from coding to teaching machines*. Massachusetts Institute of Technology.

Druga, S., Williams, R., Breazeal, C., & Resnick, M. (2017). "Hey Google is it OK if I eat you?" *Proceedings of the 2017 Conference on Interaction Design and Children - IDC '17*, 595–600. https://doi.org/10.1145/3078072.3084330

Druga, S., Williams, R., Park, H. W., & Breazeal, C. (2018). How Smart Are the Smart Toys?: Children and Parents' Agent Interaction and Intelligence Attribution. *Proceedings of the 17th ACM Conference on Interaction Design and Children*, 231–240. https://doi.org/10.1145/3202185.3202741

Ehlers, K.-H. (2009). Anrede in der Grundschule. *Zeitschrift Für Germanistische Linguistik*, *37*(2), 315–338.

Eiser, C., & Twamley, S. (1999). Talking to Children About Health and Illness. In M. Murray & K. Chamberlain (Eds.), *Qualitative Health Psychology: Theories and Methods* (pp. 133–147). London, England, UK: Sage Publications.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On Seeing Human: A Three-Factor Theory of Anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1–4. https://doi.org/10.11648/j.ajtas.20160501.11

Field, A. (2009). *Discovering Statistics Using SPSS: and sex and drugs and rock "n" roll* (3rd ed.). London, UK: Sage Publications.

Flewitt, R. (2014). Interviews. In A. Clark, R. Flewitt, M. Hammersley, & M. Robb (Eds.), *Understanding research with children and young people* (pp. 136–153). London, England, UK: Sage Publications.

Flick, U. (2014). *An Introduction to Qualitative Research* (5th ed.). Los Angeles, CA, US: Sage Publications.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Gardner, M. K. (2011). Theories of Intelligence. In E. A. Bray & T. J. Kehle (Eds.), *The Oxford Handbook of School Psychology* (pp. 79–100). https://doi.org/10.1093/oxfordhb/9780195369809.013.0035

Gaudiello, I., Lefort, S., & Zibetti, E. (2015). The ontological and functional status of robots: How firm our representations are? *Computers in Human Behavior*, *50*, 259–273. https://doi.org/10.1016/j.chb.2015.03.060

Gelman, S. A. (1988). Children's Expectations Concerning Natural Kind Categories. *Human Development*, *24*(1), 28–34. https://doi.org/10.1159/000273201

Gelman, S. A. (2013). Concepts in Development. In P. D. Zelazo (Ed.), *The Oxford handbook of developmental psychology (Vol 1): Body and mind.* (Vol. 1, pp. 542–563). https://doi.org/10.1093/oxfordhb/9780199958450.013.0019

Gelman, S. A., & Frazier, B. N. (2012). Development of Thinking in Children. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford Handbook of Thinking and Reasoning, Vol. 1: Body and Mind* (pp. 513–528). https://doi.org/10.1093/oxfordhb/9780199734689.013.0026

Given, L. M. (2008). Rigor in Qualitative Research. In N. J. Salkind (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods (Vols. 1-0)*. https://doi.org/10.4135/9781412963909 NV - 0

Glaser, B. G. (2002). Constructivist Grounded Theory? *Forum: Qualitative Social Research*, *3*(3).

Gopnik, A. (2019). AI Versus Four-Year-Olds. In J. Brockman (Ed.), *Possible Minds: Twenty-Five Ways of Looking at AI* (pp. 219–230). New York, NY, US: Penguin Press.

Gopnik, A., Griffiths, T. L., & Lucas, C. G. (2015). When Younger Learners Can Be Better (or at Least More Open-Minded) Than Older Ones. *Current Directions in Psychological Science*, *24*(2), 87–92. https://doi.org/10.1177/0963721414556653

Gopnik, A., O'Grady, S., Lucas, C. G., Griffiths, T. L., Wente, A., Bridgers, S., … Dahl, R. E. (2017). Changes in Cognitive Flexibility and Hypothesis Search Across Human Life History from Childhood to Adolescence to Adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(30), 7892–7899. https://doi.org/10.1073/pnas.1700811114

Gray, M., & Suri, S. (2017). The Humans Working Behind the AI Curtain. Retrieved April 16, 2019, from Harvard Business Review website: https://hbr.org/2017/01/the-humans-working-behind-the-ai-curtain

Greene, J. C. (2007). *Mixed Methods in Social Inquiry*. San Francisco, CA, US: Jossey-Bass.

Greene, J. C., & Hall, J. N. (2010). Dialectics and Pragmatism: Being of Consequence. In A. Tashakkori & C. Teddlie (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research* (2nd ed., pp. 119–144). https://doi.org/10.4135/9781506335193

Griffin, E. A. (1994). *A first look at communication theory* (2nd ed.). London, England, UK: McGraw-Hill.

Hammersley, M. (2011). *Methodology: Who Needs It?* London, England, UK: Sage Publications.

Hannay, J., Dudley, R., Milan, S., Kellogg Leibovitz, P., & Rodino, V. L. (2016). Qualitative Research With A "Double Life": A Mixed Methods Approach To Research and Advocacy With Adolescents. In J. Prior & J. C. Herwegen (Eds.), *Practical Research with Children* (pp. 145–165). London, England, UK: Routledge.

Harari, Y. N. (2018). *21 lessons for the 21st century*. London, England, UK: Jonathan Cape.

Harnad, S. (1991). Other Bodies, Other Minds: A Machine Incarnation of an Old Philosophical Problem. *Minds and Machines*, *1*(1), 43–54.

Hempel, J. (2015). Siri and Cortana Sound Like Ladies Because of Sexism. Retrieved March 18, 2019, from WIRED website: https://www.wired.com/2015/10/why-siri-cortana-voice-interfaces-sound-female-sexism/

Hennessy, E., & Heary, C. (2005). Exploring Children's Views Through Focus Groups. In S. Greene & D. Hogan (Eds.), *Researching Children's Experience: Methods and Approaches* (pp. 236–252). London, England, UK: Sage Publications.

Holloway, I., & Todres, L. (2003). The Status of method: flexibility, consistency and coherence. *Qualitative Research*, *3*(3), 345–357.

Holton, J. A. (2007). The Coding Process and Its Challenges. In A. Bryant & K. Charmaz (Eds.), *The SAGE Handbook of Grounded Theory* (pp. 265–289). Sage Publications.

Huitt, W., & Hummel, J. (2003). Piaget's Theory of Cognitive Development. *Educational Psychology Interactive*, *3*(2), 1–5.

Janesick, V. J. (1998). The Dance of Qualitative Research Design: Metaphor, Methodolatry, and Meaning. In N. K. Denzin & Y. S. Lincoln (Eds.), *Strategies of Qualitative Inquiry* (pp. 35–55). London, England, UK: Sage Publications.

Jick, T. D. (1979). Mixing Qualitative and Quantitative Methods: Triangulation in Action. *Administrative Science Quarterly*, *24*(4), 602–611. https://doi.org/10.2307/2392366

Jipson, J. L., & Gelman, S. A. (2007). Robots and Rodents: Children's Inferences About Living and Nonliving Kinds. *Child Development*, *78*(6), 1675 – 1688. https://doi.org/10.1016/0045-6535(89)90475-X

Jipson, J. L., Gülgöz, S., & Gelman, S. A. (2016). Parent-child conversations regarding the ontological status of a robotic dog. *Cognitive Development*, *39*, 21–35. https://doi.org/10.1016/j.cogdev.2016.03.001

Kahn, P. H., Friedman, B., Pérez-Granados, D. R., & Freier, N. G. (2006). Robotic pets in the lives of preschool children. *Interaction Studies*, *7*(3), 405–436.

Kahn, P. H., Gary, H. E., & Shen, S. (2013). Children's Social Relationships With Current and Near-Future Robots. *Child Development Perspectives*, *7*(1), 32–37. https://doi.org/10.1111/cdep.12011

Kahn, P. H. J., Reichert, A. L., Gary, H. E., Kanda, T., Ishiguro, H., Shen, S., Gill, B. (2011). The New Ontological Category Hypothesis in Human- Robot Interaction. *HRI'11 - 6th Annual Conference for Basic and Applied Human-Robot Interaction Research*, 159–160. Lausanne, Switzerland.

Kahn, P. H., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., Gill, B. T., Shen, S. (2012). "Robovie, you'll have to go into the closet now": Children's social and moral relationships with a humanoid robot. *Developmental Psychology*, *48*(2), 303–314. https://doi.org/10.1037/a0027033

Karelis, C. (1986). Reflections on the Turing Test. *Journal for the Theory of Social Behaviour*, *16*(2), 161–172. https://doi.org/10.1111/j.1468-5914.1986.tb00073.x

Karrer, K., Glaser, C., Clemens, C., & Bruder, C. (2009). Technikaffinität erfassen: Der Fragebogen TA-EG. *Der Mensch Im Mittelpunkt Technischer Systeme : 8. Berliner Werkstatt Mensch-Maschine-Systeme, 7. Bis 9. Oktober 2009*, 196–201. Düsseldorf, North Rhine-Westphalia, Germany: VDI Verlag Düsseldorf.

Kelle, U., Kühlberger, C., & Bernhard, R. (2019). How to use Mixed Methods and Triangulation designs - an introduction to research in History Education. In R. Bernhard (Ed.), *Mixed Methods/Triangulation in History Education Research* (Special Ed). Open Access.

Kinsella, B., & Mutchler, A. (2019). *Smart Speaker Consumer Adoption Report*.

Kissinger, H. A. (2018). How the Enlightenment Ends. Retrieved March 1, 2019, from The Atlantic website: https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/

Kortesluoma, R., Hentinen, M., & Nikkonen, M. (2003). Conducting A Qualitative Child Interview: Methodological Considerations. *Journal of Advanced Nursing*, *42*(5), 434–441. https://doi.org/10.1046/j.1365-2648.2003.02643.x

Kranzberg, M. (1986). Technology and History: "Kranzberg's Laws." *Technology and Culture*, *27*(3), 544–560.

Krueger, R. A., & Casey, M. A. (2000). *Focus groups: A Practical Guide for Applied Research* (3rd ed.). Thousand Oaks, CA, US: Sage Publications.

Ku, H., Choi, J. J., Lee, S., Jang, S., & Do, W. (2018). Designing Shelly, a Robot Capable of Assessing and Restraining Children's Robot Abusing Behaviors. *ACM/IEEE International Conference on Human-Robot Interaction*, 161–162. https://doi.org/10.1145/3173386.3176973

Kurzweil, R. (1990). *The Age of Intelligent Machines*. Cambridge, MA, US: MIT Press.

Lapadat, J. C., & Lindsay, A. C. (1999). Transcription in Research and Practice: From Standardization of Technique to Interpretive Positionings. *Qualitative Inquiry*, *5*(1), 64–86. https://doi.org/10.1177/107780049900500104

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lee, K.-F. (2018). *AI superpowers: China, Silicon Valley, and the new world order*. Boston, MA, US: Houghton Mifflin Harcourt.

Leech, N. L., & Onwuegbuzie, A. J. (2002). A Call for Greater Use of Nonparametric Statistics. *Annual Meeting of the Mid-South Educational Research Association*. Chattanooga, TN, US: ERIC.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. Beverly Hills, CA, US: Sage Publications.

Lovato, S., & Piper, A. M. (2015). "Siri, is this you?": Understanding Young Children's Interactions with Voice Input Systems. *Proceedings of the 14th International Conference on Interaction Design and Children - IDC '15*, 335–338. https://doi.org/10.1145/2771839.2771910

Mackenzie, N., & Knipe, S. (2006). Research Dilemmas: Paradigms, Methods and Methodology. *Issues In Educational Research*, *16*(2), 193–205.

Mao, J., Gan, C., Kohli, P., Tenenbaum, J. B., & Wu, J. (2019). The Neuro-Symbolic Concept Learner: Interpreting Scenes, Words, And Sentences From Natural Supervision. *ICLR 2019 - Seventh International Conference on Learning Representations*. New Orleans, LA, US.

Martinich, A. P., & Stroll, A. (2017). epistemology. In *Encyclopaedia Britannica Online*. Encyclopaedia Britannica, Inc.

Mauthner, M. (1997). Methodological Aspects of Collecting Data from Children: Lessons from three Research Projects. *Children & Society*, *11*(1), 16–28. https://doi.org/10.1111/j.1099-0860.1997.tb00003.x

Mayall, B. (1994). *Negotiating Health: Children at home and Primary School*. London, England, UK: Cassell.

Mays, N., & Pope, C. (1995). Rigour And Qualitative Research. *British Medical Journal*, *311*(6997), 109–112. https://doi.org/10.17818/NM/2016/4.3

Mazhelis, O., Luoma, E., & Warma, H. (2012). Defining an Internet-of-Things Ecosystem. In S. Andreev, S. Balandin, & Y. Koucheryavy (Eds.), *Proceedings of the Conference on Internet of Things and Smart Spaces, & the International Conference on Next Generation Wired/Wireless Networking - ruSMART 2012 & NEW2AN 2012* (pp. 1–14). Berlin: Springer.

Melrose, S. (2010). Naturalistic Generalization. In A. J. Mills, G. Durepos, & E. Wiebe (Eds.), *Encyclopedia of Case Study Research* (pp. 600–601). Thousand Oaks, CA, US: Sage Publications.

Moor, J. H. (1976). An Analysis of the Turing Test. *Philosophical Studies*, *30*(4), 249–257.

Morgan, M., Gibbs, S., Maxwell, K., & Britten, N. (2002). Hearing Children's Voices: Methodological Issues in Conducting Focus Groups with Children Aged 7-11 Years. *Qualitative Research*, *2*(1), 5–20.

Mumford, L. (2010). *Technics and Civilization* (L. Winner, Ed.). London, England, UK: University of Chicago Press.

Nass, C. I., & Brave, S. (2005). *Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship*. Cambridge, MA, US: MIT Press.

O'Cathain, A. (2010). Assessing the Quality of Mixed Methods Research: Toward a Comprehensive Framework. In A. Tashakkori & C. Teddlie (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research* (2nd ed., pp. 531–556). Thousand Oaks, CA, US.

O'Dell, L., & Brownlow, C. (2016). Ethical Practice in Research About Children. In J. Prior & J. Van Herwegen (Eds.), *Practical Research with Children* (pp. 285–300). London, England, UK: Routledge.

OECD. (2014). Indicator D1: How much time do students spend in the classroom? In *Education at a Glance 2014: OECD Indicator*. https://doi.org/http://dx.doi.org/10.1787/888933119530

Onwuegbuzie, A. J., & Combs, J. P. (2010). Emergent Data Analysis Techniques in Mixed Methods Research: A Synthesis. In A. Tashakkori & C. Teddlie (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research* (2nd ed., pp. 397–430). Thousand Oaks, CA, US: Sage Publications.

Pearl, J. (2019). The Limitations of Opaque Learning Machines. In J. Brockman (Ed.), *Possible Minds: Twenty-Five Ways of Looking at AI* (pp. 13–19). New York, NY, US: Penguin Press.

Peterson, R. A. (1994). A Meta-Analysis of Cronbach's Coefficient Alpha. *Journal of Consumer Research*, *21*(2), 381–391.

Porcheron, M., Fischer, J. E., Reeves, S., & Sharples, S. (2018). Voice Interfaces in Everyday Life. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1–12. https://doi.org/10.1145/3173574.3174214

Postbank. (2019). *Postbank Digitalstudie 2019: Untersuchung zur Internet und Mobilnutzung der Bevölkerung*. Bonn, Germany.

Prior, J. (2016). The Use Of Semi-Structured Interviews With Young Children. In J. Prior & J. Van Herwegen (Eds.), *Practical Research with Children* (pp. 109–126). London, England, UK: Routledge.

Purington, A., Taft, J. G., Sannon, S., Bazarova, N. N., & Taylor, S. H. (2017). "Alexa is my new BFF": Social roles, user satisfaction, and personification of the Amazon Echo. *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*, *Part F1276*, 2853–2859. https://doi.org/10.1002/ece3.1301

PwC. (2018). *Global Top 100 Companies by Market Capitalisation - 31 March 2018 update*. London, England, UK.

Quantcast. (2019). *Quantcast Audience Insights: Sprachassistentin Alexa spricht vor allem Familien mit Kindern an*. Munich, Bavaria, Germany.

Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., … Wellman, M. (2019). Machine behaviour. *Nature*, *568*(7753), 477–486. https://doi.org/10.1038/s41586-019-1138-y

Ramakrishnan, V. (2019). Will Computers Become Our Overlords? In J. Brockman (Ed.), *Possible Minds: Twenty-Five Ways of Looking at AI* (pp. 181–191). New York, NY, US: Penguin Press.

Reeves, B., & Nass, C. I. (1998). *The media equation: How people treat computers, television, and new media like real people and places*. Stanford, CA, US: CSLI Publications.

Resnik, D. B. (2015). *What is Ethics in Research & Why is it Important?*

Rice, M., Bunker, K. D., Kang, D.-H., Howell, C. C., & Weaver, M. (2007). Accessing and Recruiting Children for Research in Schools. *Western Journal of Nursing Research*, *29*(4), 501–514. https://doi.org/10.1177/0193945906296549

Riessman, C. K. (1993). *Narrative Analysis*. London, England, UK: Sage Publications.

Robinson, O. C. (2014). Sampling in Interview-Based Qualitative Research: A Theoretical and Practical Guide. *Qualitative Research in Psychology*, *11*(1), 25–41. https://doi.org/10.1080/14780887.2013.801543

Rodriguez, S. (2019). Facebook building voice assistant to rival Amazon Alexa and Apple Siri. Retrieved April 26, 2019, from CNBC website: https://www.cnbc.com/2019/04/17/facebook-building-voice-assistant-to-rival-amazon-alexa-and-apple-siri.html

Ross, C. (2019). New voices at patients' bedsides: Amazon, Google, Microsoft, and Apple. Retrieved February 28, 2019, from STAT website: https://www.statnews.com/2019/02/06/voice-assistants-at-bedside-patient-care/?utm_campaign=the_download.unpaid.engagement&utm_source=hs_email&utm_medium=email&utm_content=69643770&_hsenc=p2ANqtz--nOHcxWiobOCFlnUebi0I5Zc_2HUuv1GHjlPLxaPehTUojtzL-mYVdu0NjNz

Samuels, A., & Taylor, M. (1994). Children's ability to distinguish fantasy events from real-life events. *British Journal of Developmental Psychology*, *12*(4), 417–427. https://doi.org/10.1111/j.2044-835X.1994.tb00644.x

Saygin, A. P., Cicekli, I., & Akman, V. (2001). Turing Test: 50 Years Later. *Minds and Machines*, *10*, 463–518.

Scaife, M., & Duuren, M. (1995). Do computers have brains? What children believe about intelligent artifacts. *British Journal of Developmental Psychology*, *13*(4), 367–377. https://doi.org/10.1111/j.2044-835X.1995.tb00686.x

Schwab, K. (2017). *The fourth industrial revolution*. Geneva, Switzerland: World Economic Forum.

Searle, J. R. (1990). Is the Brain's Mind a Computer Program. *Scientific American*, *262*(1), 26–31.

Searle, J. R. (2009). Chinese room argument. *Scholarpedia*, *4*(8), 3100.

Severson, R. L., & Carlson, S. M. (2010). Behaving as or behaving as if? Children's conceptions of personified robots and the emergence of a new ontological category. *Neural Networks*, *23*(8–9), 1099–1103. https://doi.org/10.1016/j.neunet.2010.08.014

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Belmont, CA, US: Wadsworth Cengage Learning.

Shavlik, J. W. (1994). Combining Symbolic And Neural Learning - Extended Abstract. *Machine Learning*, *14*(3), 321–331.

Shaw, T., Cross, D., Thomas, L. T., & Zubrick, S. R. (2015). Bias in Student Survey Findings from Active Parental Consent Procedures. *British Educational Research Journal*, *41*(2), 229–243. https://doi.org/10.1002/berj.3137

Shiomi, M., Kanda, T., Howley, I., Hayashi, K., & Hagita, N. (2015). Can a Social Robot Stimulate Science Curiosity in Classrooms? *International Journal of Social Robotics*, *7*(5), 641–652. https://doi.org/10.1007/s12369-015-0303-1

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., … Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, *362*(6419), 1140–1144. https://doi.org/10.1126/science.aar6404

Simons, P. M. (2009). ontology. In *Encyclopaedia Britannica Online*. Encyclopaedia Britannica, Inc.

Small, C. E., Nigrin, D., Churchwell, K., & Brownstein, J. (2018). What Will Health Care Look Like Once Smart Speakers Are Everywhere? Retrieved February 28, 2019, from Harvard Business Review website: https://hbr.org/2018/03/what-will-health-care-look-like-once-smart-speakers-are-everywhere

Sue, V. M., & Ritter, L. A. (2012). *Conducting Oline Surveys* (2nd ed.). Los Angeles, CA, US: Sage Publications.

Tanaka, F., Cicourel, A., & Movellan, J. R. (2007). Socialization between toddlers and robots at an early childhood education center. *Proceedings of the National Academy of Sciences*, *104*(46), 17954–17958. https://doi.org/10.1073/pnas.0707769104

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA, US: Sage Publications.

Tashakkori, A., & Teddlie, C. (2008). Quality of Inferences in Mixed Methods Research: Calling for an Integrative Framework. In M. M. Bergman (Ed.), *Advances in Mixed Methods Research: Theories and Applications* (pp. 101–119). Los Angeles, CA, US: SAGE.

Tashakkori, A., & Teddlie, C. (2010). Epilogue: Current Developments and Emerging Trends in Integrated Research Methodology. In A. Tashakkori & C. Teddlie (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research* (2nd ed., pp. 803–826). https://doi.org/10.4135/9781506335193

Thompson, N. (2019). An AI Pioneer Explains The Evolution of Neural Networks. Retrieved May 18, 2019, from WIRED website: https://www.wired.com/story/ai-pioneer-explains-evolution-neural-networks/?utm_campaign=the_algorithm.unpaid.engagement&utm_source=hs_email&utm_medium=email&utm_content=72659496&_hsenc=p2ANqtz-8bs2iv7YTA5GAstR1JvuB4Qond4xhyEHoB6UP_21cUs_XWPwtnnrXu1WCGM6TC

Tobin, G. A., & Begley, C. M. (2004). Methodological rigour within a qualitative framework. *Journal of Advanced Nursing*, *48*(4), 388–396.

Turing, A. (1950). Computing machinery and intelligence. *Mind*, *49*, 433–460.

Turkle, S. (2005). *The Second Self: Computers and the Human Spirit* (Twentieth). London, England, UK: The MIT Press.

UNESCO. (2019). I'd blush if I could. In *EQUALS Skills Coalition*.

Van Duuren, M., Dossett, B., & Robinson, D. (1998). Gauging Children's Understanding of Artificially Intelligent Objects: A Presentation of "Counterfactuals." *International Journal of Behavioral Development*, *22*(4), 871–889. https://doi.org/ 10.1080/016502598384207

Velik, R., & Bruckner, D. (2008). Neuro-Symbolic Networks: Introduction to a New Information Processing Principle. *IEEE 2008 - 6th IEEE International Conference on Industrial Informatics*, 1042–1047. https://doi.org/10.1109/INDIN.2008.4618256

Vollmer, A. L., Read, R., Trippas, D., & Belpaeme, T. (2018). Children conform, adults resist: A robot group induced peer pressure on normative social conformity. *Science Robotics*, *3*(21). https://doi.org/10.1126/scirobotics.aat7111

Watt, S. (1996). Naive Psychology and the Inverted Turing Test. *Psycoloquy*, *7*(14).

Waytz, A., Cacioppo, J., & Epley, N. (2014). Who Sees Human? The Stability and Importance of Individual Differences in Anthropomorphism. *Perspectives on Psychological Science*, *5*(3), 219–232. https://doi.org/10.1177/ 1745691610369336.

Webster, F. (2006). *Theories of the Information Society* (3rd ed.). https://doi.org/10.4324/9781315867854

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief. *Child Development*, *72*(3), 655–684. https://doi.org/10.1111/1467-8624.00304

Wellman, H. M., & Gelman, S. A. (1992). Cognitive Development: Foundational Theories of Core Domains. *Annual Review Psychology*, *43*, 337–375.

Wheeler, M. (2008). God's Machines: Descartes on the Mechanization of Mind. In P. Husbands, O. Holland, & M. Wheeler (Eds.), *The Mechanical Mind in History* (pp. 307–330). Cambridge, MA, US: Oxford University Press.

Wiggers, K. (2019). Google Assistant will soon be on 1 billion devices, but still can't speak like John Legend. Retrieved January 14, 2019, from VentureBeat website: https://venturebeat.com/2019/01/07/google-assistant-will-soon-be-on-1-billion-devices/

Wilczek, F. (2019). The Unity of Intelligence. In J. Brockman (Ed.), *Possible Minds: Twenty-Five Ways of Looking at AI* (pp. 66–75). New York, NY, US: Penguin Press.

Winchatz, M. R. (2001). Social Meanings in German Interactions: An Ethnographic Analysis of the Second-Person Pronoun Sie. *Research on Language & Social Interaction*, *34*(3), 337–369. https://doi.org/10.1207/S15327973RLSI34-3_3

Wood, L., Dautenhahn, K., Robins, B., & Zaraki, A. (2017). Developing child-robot interaction scenarios with a humanoid robot to assist children with autism in developing visual perspective taking skills. *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 1055–1060. https://doi.org/10.1109/ROMAN.2017.8172434

Woolley, J. D. (1997). Thinking about Fantasy: Are Children Fundamentally Different Thinkers and Believers from Adults? *Child Development*, *68*(6), 991–1011. https://doi.org/10.2307/1132282

Yarosh, S., Thompson, S., Watson, K., Chase, A., Senthilkumar, A., Yuan, Y., & Brush, A. J. B. (2018). Children Asking Questions: Speech Interface Reformulations and Personification Preferences. *Proceedings of the 17th ACM Conference on Interaction Design and Children*, 300–312. https://doi.org/10.1145/3202185.3202207

Zelazo, P. D. (2013). Developmental Psychology: A New Synthesis. In P. D. Zelazo (Ed.), *The Oxford Handbook of Developmental Psychology, Vol. 1: Body and Mind* (pp. 3–12). https://doi.org/10.4135/9781446214633

Złotowski, J., Proudfoot, D., Yogeeswaran, K., & Bartneck, C. (2015). Anthropomorphism: Opportunities and Challenges in Human – Robot Interaction. *International Journal of Social Robotics*, *7*, 347–360. https://doi.org/10.1007/s12369-014-0267-6

Appendix

Appendix I. Excursus: Conceptual parallels between the nature of HI and MI

This excursus provides a brief introduction to Gopnik's (2019) recent essay on the scientific hypothesis that there might be substantial HI-MI parallels, namely between the inferential processes that drive human cognitive development on the one hand, and some of the most common approaches that are currently used in the realm of MI for inductive data-driven inference on the other hand.

In a nutshell, Gopnik (2019) argues that today's most plausible explanation for the mechanics of humans' cognitive development is indeed that "our brains perform computations on the concrete, particular, messy data arriving at our senses, and those computations yield accurate representations of the world" (p.222). This argument itself is based on the philosophically grounded assumption that there are two distinct ways for the inductive inference of conceptual patterns from reality, referred to as the *'bottom-up*' and '*top-down*' approach of learning (Gopnik, 2019).

According to the bottom-up approach, conceptual patterns of reality emerge from the vast array of raw sensory impressions and only depend on the capacity to identify structures and to derive abstractions from them (Gopnik, 2019). In contrast, the top-down argument states that any raw input of sensory impressions is already approached with an a priori understanding of reality that leverages how these new inputs are exploited in order to refine existing representations (Gopnik, 2019). The major difference between these two approaches is that explorative bottom-up inference requires large amounts of sensory input in order to derive original abstractions, but in turn the obtained understanding can be very *archetypical* and *genuine,* since it is completely driven by raw sensory inputs (Gopnik, 2019). In contrast, top-down inference does not necessarily require such large amounts of sensory input; instead a sufficient prior conceptualisation of reality is needed, which can leverage but also bias the inferential process of refining existing conceptualisations according to the new input, therefore limiting the degrees of originality or genuineness (Gopnik, 2019).

When it comes to humans' cognitive learning trajectories, Gopnik (2019) further argues that children seem to rely on a synergy of both approaches depending on their developmental stage, because even infants can extend their conceptual understanding of reality based on a few cases of sensory input (i.e. top-down), but they can also form very original and genuine concepts (i.e. bottom-up) based on a vast array of raw sensory impressions (Gopnik, Griffiths, & Lucas, 2015; Gopnik et al., 2017).

Other authors further argue that this unique dualism of humans' cognitive learning trajectories has not only enabled our species' evolutionary success in the past, but it also constitutes the major difference between human intelligence and MI, at least up till now (cf. Pearl, 2019). Indeed, when it comes to machines, the most common approaches that are currently used in the realm of MI for inductive data-driven inference either rely on the bottom-up approach, meaning that algorithms explore huge amounts of data in order to identify patterns, or on the top-down approach, meaning that algorithms are a priori endowed with concrete concepts that guide their inference from the data (Gopnik, 2019).

Such MI-driven *bottom-up* approaches are typically referred to as *connectionist systems,* while *top-down* approaches are referred to as *symbolic systems* (Gopnik, 2019). Within the connectionist domain, today's most common technological applications are *deep-learning* or *reinforcement learning* (or a combination of both) based on artificial neural networks, while the symbolic domain mostly uses *generative Bayesian models* (Gopnik, 2019). Although their synergetic effects seem obvious, these two approaches have developed rather separately ever since the emergence of MI in the 1950s (Shavlik, 1994; Velik & Bruckner, 2008), and only recently joint efforts have achieved first breakthroughs (e.g. Mao, Gan, Kohli, Tenenbaum, & Wu, 2019).

Hence, Gopnik's (2019) argument concludes that HI's superiority in terms of inductive inference relies on MI's current limitation to combine both approaches within the same computational entity, further implying that most of today's MI-applications struggle with the implicit limitations that are associated with the approach they rely on (Gopnik, 2019; Velik & Bruckner, 2008). For instance, even the most proficient algorithms that follow the bottom-up approach need huge amounts of data in order to extend their conceptual horizon (e.g. identifying a new kind of animal on a picture), while a child can achieve the same based on a few examples only. However, it should also be mentioned that the data-driven bottom-up approach has enabled MI to surpass HI in domains where we used to have more than a thousand years of experience, for instance when it comes to board games, such as 'Go' (cf. Anderson, 2019; Silver et al., 2018).

Notes. Table provides a brief introduction to Gopnik's (2019) recent essay on the scientific hypothesis that there might be substantial HI-MI parallels. This discussion of Gopnik's (2019) hypothesis is replenished with additional references from the literature.

Appendix II. Detailed description of research procedure

Part	Title	Time Period	Description
(I)	Desk research	January to March	• Reviewing relevant empirical, theorical, and methodological literature for the preparation of the study
(II)	CUREC	Early March	• Submission of documents which are required by CUREC at University of Oxford to conduct research that involves human participants • Documents submitted to DREC at the Department of Education • DREC granted approval in late March 2019
(III)	Preparation	March to April	• Contacting headteachers of Chandler's primary schools in order to sound out general interest to participate in the study • After internal consultations with teaching staff and parent representatives both headteachers confirmed by mid-March that their institutions would be willing to participate • On request of both headteachers one-page information letters were distributed to all parents before Easter Break • In information letters the SI introduced himself as a local and as a former student of school 1, explained the topic of the study, its importance, the means of data collection and additional points of emphasis (i.e. reassuring parents that the study was not related to the children's school performance in any way and emphasising that everyone was very welcome to participate, regardless of the personal usage of or attitude towards technology) • Printed letters had to be distributed through class-teachers because due to data protection regulations schools did not use email addresses to communicate with parents
(IV)	Translation and piloting	Mid-April	• During two-week long Easter Break all drafts and materials were translated by the SI (English ⟶ German) • Parental survey as well as focus groups were piloted with a small group of volunteering individuals who were recruited within the SI's broader circle of family and friends (further details in section 2.3)
(V)	In-class familiarisation	Late April to early May	• After Easter break in-class familiarisation began in school 1 (in-class familiarisation in school 2 started one week later and followed the same procedure as described below) • During this phase SI spent one week in several classrooms across grades 1 to 4 (before each lesson, teachers consented verbally to presence of SI in classroom) • Depending on the teacher as well as the nature of the lesson personal engagement with children was possible (e.g. during in-class activities and open work periods) but handled in a rather passive and restricted manner in order to avoid distractions from the lesson • Data collection during this phase was limited to SI's observations and respective field notes (i.e. no personal data was collected)
(VI)	Parental survey	Mid-May to late May	• Towards the end of in-class familiarisation printed survey invitations were distributed through class-teachers to all parents

			• Invitations provided brief summary of previous information letter and additional details regarding survey (e.g. voluntary nature, content, length, anonymity, deadline) • Parents were able to access online questionnaire via an anonymous QR code or an anonymous URL • On the second page of the invitation focus groups were introduced and briefly explained (e.g. voluntary nature, content, dates, importance of dual participation in both focus groups, collection of personal data through audio-visual recordings, confidentiality of data, linkage between survey and focus group data based on key codes, parental consent, no compensations for participation) • Consent forms were attached to survey invitations with additional instruction that they had to be returned to class teachers before the date of the first focus group session • After distribution of invitations online questionnaire remained active for two weeks • Secretary's office and class teachers agreed to collect signed consent forms until predefined threshold was met based on a first-come-first-served basis and to communicate on behalf of the SI that the reason behind exclusion was entirely due to external resource constraints and not of discriminatory nature
(VII)	Focus groups	Late May to mid-June	• Focus groups took place on school premises in order to avoid logistical complications for participating individuals and to build on classroom environments children were already familiar with • For safeguarding reasons members of the school's teaching staff were available at all times but without actively engaging in the data collection procedures • All sessions were scheduled directly after school because during the internal considerations parent representatives requested that the study should not interfere with the official timetable (availability of classrooms was also much higher after school therefore allowing for more flexibility regarding the selection of focus group locations) • Both sessions were conducted by the SI in the same audio-visually recorded classrooms and lasted for approximately 90 minutes each including all organisational procedures (e.g. verifying consent forms, distribution of name tags, briefing and debriefing of participants) • There was a one-week time gap between first and the second focus group session in order to allow for an immediate transcription
(VIII)	Conclusion	Mid-June	• By mid-June data collection was completed in school 1 and school 2 • An informal debriefing was scheduled with both headteachers in order to thank them for their support and to indicate the time line for the remaining research procedure as well as the availability of results • Data analysis began in June

Notes. Table provides detailed chronological overview of the research procedure. This overview depicts all major steps related to the collection of primary data that took place after the initial desk research (January – March) and before the data analysis and write-up (June – August). All time points refer to the calendar year 2019.

Appendix III. Checklist to use online questionnaires in a survey design

List No.	Checklist Item	Arguments in the context of the given study	Status
1	*Is the desired sample size and its geographical dispersion high enough to make the use of an online questionnaire feasible?*	The desired sample size for the parental survey was supposed to reach a maximum ($n \rightarrow 600$) by distributing survey invitations among all parents whose children attended Chandler's primary schools (2.2.1). Although parents resided in the same municipality, alternative means of data collection such as telephone or face-to-face interviews were feasible due to resource constraints.	Fulfilled
2	*Does the implementation of an online questionnaire – including its programming, piloting, distribution and response period – fit within the study's time constraints?*	The study's time constraints allowed for an appropriate implementation of the online questionnaire, including its programming, piloting, distribution and a response period of two weeks (2.2.2).	Fulfilled
3	*Is the online questionnaire seeking personal or even sensitive information from participants?*	The online questionnaire was completely anonymous (i.e. no personal information) and did not include any sensitive items (2.3.2).	Fulfilled
4	*Can the target population be reached through the technical means of an online questionnaire?*	The target population of the parental survey were parents and guardians of normally developing primary school children between the age of 6 and 10 in industrialised countries (2.2.1). Therefore, assuming the general prevalence of computer or mobile devices, internet access as well as minimal levels of digital literacy seemed to be reasonable in this context.	Fulfilled
5	*Can the online questionnaire be distributed appropriately?*	Although parents could not be reached via email, paper-based communication allowed for the distribution of pre-notifications as well as survey invitations among all parents in Chandler's primary school system (2.2.2).	Fulfilled

6	*Is non-random sampling acceptable (since online questionnaires typically do not allow for a true random selection)?*	Since the generalisation of claims was not a primary concern of this study, the absence of a non-random sampling was considered to be acceptable.	Fulfilled
7	*Can the online questionnaire make use of media or other interactive contents?*	The implemented online questionnaire made strong use of conditional items (e.g. contingency or follow-up questions), meaning that throughout the questionnaire some items were only shown to respondents or skipped if certain conditions in their previous responses have been met (**Appendix II**). This interactive feature had the strong advantage that only relevant items are shown to respondents.	Fulfilled
8	*Can the online questionnaire be implemented appropriately by the researcher?*	Due to prior experience with online survey tools and the general user-friendliness of the online experience management platform 'Qualtrics', the SI was able to implement the online questionnaire appropriately.	Fulfilled

Source. Checklist based on Sue & Ritter (2012).

Appendix IV. Detailed structure of online questionnaire used in parental survey

Item No.	Item content	Data entry
1	[Introduction] Dear Parents, dear guardians **Welcome** to the survey platform of the **Department of Education** [hyperlink: www.education.ox.ac.uk] at the **University of Oxford** [hyperlink: http://www.ox.ac.uk]. **Thank you** already in advance that you have decided to take part in this anonymous survey! Your participation will help us to gain a broad overview on how **digital technologies & media** affect the **daily life** of **parents and guardians** as well as their **children**. The survey consists of **three parts** and participation should take **20 minutes** on average. Most of the statements simply require you to **click** or **select** one of the response options, while a few statements are **open-entry** questions (this means that statements have to be typed in manually). **Let's go!** [emojis: 💪🎊🎉] [Name student investigator] & [Name principal investigator]	[None]
	Page break	
2	**Section I: Digital Technologies & media in daily life** [emojis: 📱💻🎧] In the following section we would like to know how digital technologies & media affect **your personal daily life.** **What do we mean by digital technologies & media?** • **Devices:** Smartphones, tablets, laptops, computers, smart TVs, Apple TV, Amazon Fire TV, smart watches, voice assistants and so on • **Communication & social media:** WhatsApp, iMessage, emails, Facebook, Facebook Messenger, Instagram, Snapchat and so on • **Entertainment:** YouTube, computer- and app games, music and audiobook streaming (Spotify, Prime Music and so on), video streaming (Netflix, Prime Video and so on) • **Organisation, hobbies, and practical functions:** Online shopping, surfing on the internet in general, online banking, fitness apps, smart home application and so on	[None]

	Page break	
3	Let's start with an introductory question: Do you use **digital technologies & media** in your daily life? **Comment:** This question refers to all types of privately used digital technologies & media, regardless how often or for what purposes you may use them.	□ [A] = Yes, I do use digital technologies & media in my daily life. □ [B] = No, I do not use digital technologies & media *at all* in my daily life.
	Page break	
4.1 – 4.16	[If item 3 = A] In the following you may find a list of **statements**. For each statement, please indicate **how well this statement applies to you personally or rather to what extent you agree with it**.	Likert Scales: □ [-2] = Totally disagree □ [-1] = Rather Disagree □ [0] = Neutral □ [1] = Rather Agree □ [2] = Totally agree
	I love it to use new digital technologies & media! [Positively poled item]	
	I like to go to specialist shops in order to seek advice on digital technologies & media! [Positively poled item]	
	I would have a hard time if I had to seek information on digital technologies & media without somebody's help! [Negatively poled item]	
	Using digital technologies & media increases the standard of living! [Positively poled item]	
	Using digital technologies & media cause intellectual impoverishment! [Negatively poled item]	
	Using digital technologies & media complicates lots of things! [Negatively poled item]	
	I seek information on technological trends and developments, even if I have no intention to buy something! [Positively poled item]	
	I enjoy trying out new digital technologies & media! [Positively poled item]	

	Digital technologies & media make my daily life easier! [Positively poled item]	
	Using digital technologies & media is a security and privacy threat! [Negatively poled item]	
	Digital technologies & media reduce the personal contact among humans! [Negatively poled item]	
	I know most of the functions of the digital devices and application I use! [Positively poled item]	
	I am excited when new digital technologies & media are released to the market. [Positively poled item]	
	Using digital technologies & media is causing stress! [Negatively poled item]	
	It is easy for me to learn how to use new digital technologies & media! [Positively poled item]	
	Hopefully digital technologies & media will be introduced to more areas of daily life! [Positively poled item]	

Page break

5	**How much** do you have to deal with technological trends and developments in your **daily work environment**? **Comment:** This question especially refers to how much the pursuance of your profession has been affected by technological trends and developments in recent years, or how regularly this topic plays a role in your daily work environment.	Likert Scales: □ [1] = Never □ [2] = Rarely □ [3] = Occasionally □ [4] = Often □ [5] = Very Often □ [99] = I am currently unemployed.

Page break

6	[If item 3 = B] What were or what are the **main reasons** for you **to refrain from the usage of digital technologies & media**?	[Open-entry]

Page break		
7	**Section II: Is this smart or what? Artificial Intelligence in daily life** [emojis: 💡⚙️💬] The second part of the survey is about artificial intelligence. **Important:** This part is based on very general questions that do not require any specialist knowledge! Following that, we would also like to know whether you own a digital voice assistant and how you use it in your daily life.	[None]
	Page break	
8	Let's start with another introductory question: **How familiar** are you with the term '**Artificial Intelligence'** (often abbreviated with '**A.I.'**)?	Likert Scales: □ [1] = Not at all familiar / Never heard of it □ [2] = Slightly familiar / Have heard of it a once or twice □ [3] = Somewhat familiar / Hear it from time to time □ [4] = Moderately familiar / Have read up on it a little bit □ [5] = Extremely familiar / Know a lot about it
	Page break	
9.1 – 9.5	[If item 8 = 2 or 3 or 4 or 5] In the following you may find a list of **statements** again. For each statement, please indicate **how well this statement applies to you personally or rather to what extent you agree with it**.	Likert Scales: □ [-2] = Totally disagree □ [-1] = Rather Disagree □ [0] = Neutral □ [1] = Rather Agree
	I find it convenient that a growing number of tasks and processes are becoming automated based on A.I.! [Positively poled item]	

	It is progressive that nowadays machines are able to improve themselves while carrying out certain tasks! [Positively poled item]	□ [2] = Totally agree
	I believe that today, or at some point in the future, artificial intelligence is, or will be, a great threat to society! [Negatively poled item]	
	If A.I. is able to carry out a certain task in a more accurate and reliable way compared to a human, the application of the respective technology should be considered! [Positively poled item]	
	I am worried that the future prospects of younger generations will be negatively affected by the growing automation of tasks and processes! [Negatively poled item]	

Page break

10	**Do you own devices** at home that allow you or other family members to use **digital voice assistants**? **Comment:** This question refers to digital voice assistants such as *Alexa*, *Siri*, *Google*, *Cortana* or similar applications which can be accessed through smartphones, smart speakers or laptops.	□ [A] = Yes, we do own such devices and make use of digital voice assistants. □ [B] = Yes, we do own devices at home which would allow us to use digital voice assistants, but we do not make use of them. □ [C] = No, we do not own any devices at all which would allow us to use digital voice assistants.

Page break

11	[If item 10 = A] **What digital voice assistants** do you or other family members use? **Comment:** If you own or even use multiple digital voice assistants, please select the one that you think is used the most in your household.	□ [A] = Apple's Siri □ [B] = Amazon's Alexa □ [C] = Google's Google Home

		□ [D] = Microsoft's Cortana □ [E] = Other: [open-entry]
12	[If item 10 = A] **On how many devices** do you or other family members use the digital voice assistant selected above?	□ [1] = 1 device □ [2] = 2 devices □ [3] = 3 devices □ [4] = 4 or more devices
13	[If item 10 = A] **How long have** you or other family members used the digital voice assistant selected above?	□ [1] = Less than 1 year □ [2] = Between 1 and 2 years □ [3] = Between 2 and 3 years □ [4] = More than 3 years
14	[If item 10 = A] **Where** do you or other family members use the digital voice assistant selected above? **Comment:** Multiple selections possible.	□ [A] = Dining room □ [B] = Bedroom □ [C] = Bathroom □ [D] = Children's room □ [E] = Living room □ [F] = Other locations: [open-entry]
15	[If item 10 = A] **How much** do you or other family members use the digital voice assistant for the following purposes?	Likert Scales: □ [1] = Never □ [2] = Rarely □ [3] = Occasionally □ [4] = Often □ [5] = Very Often
	Music, audiobooks, podcasts or other audio entertainment **Examples:** Spotify, Prime Music, Apple Music, audio book apps and so on	
	Games or other forms of *interactive* entertainment **Examples:** Asking funny questions, playing Bingo/Akinator/Rock-paper-scissors or similar applications	

	Daily practical assistance **Examples:** Setting timers/alarms/reminders/calendar entries or similar things, seeking weather information, online shopping and so on	
	Controlling smart home devices **Examples:** Smart plugs, light bulbs, thermostat, door bells and so on	

Page break

16	**Do people in your closer circle of family and friends** use digital voice assistants?	□ [A] = Yes, there are people in my closer circle of family and friends who use digital voice assistants. □ [B] = No, I do not know anybody within my closer circle of family and friends who uses digital voice assistants.
17	[If item 16 = A] Does **your child** have the opportunity **to engage with digital voice assistants** in your **closer circle of family and friend**s?	□ [A] = Yes, this can happen. □ [B] = No, up to my knowledge this does not happen.

Page break

18	[If item 10 = B or C] What were or what are the **main reasons** for you **to refrain from the usage of digital voice assistants**?	□ [A] = Expected utility and value is too little □ [B] = Too expensive to buy □ [C] = Concerns regarding privacy and data protection □ [D] = Did not know that there are things like “digital voice assistants”

		□ [E] = Installation and usage would be too complicated □ [F] = Other reasons: [open-entry]
Page break		
19	**Section III: Children and digital technologies & media in daily life** [emojis: 📱💻🎧] In the third and last section of this survey we would like to what extent your child uses digital technologies & media. **Comments:** • This section refers to your child that currently goes to [name of participating primary school]. • For households with multiple children who currently go to [name of participating primary school]: For our study it would be sufficient if you focus on one of the children living in your household, given that your answers to the following questions differ for your children.	[None]
Page break		
20	Let's start with some introductory question again: **How many children currently** live in your household? **Comment:** This question refers to all individuals who are below the age of 21. and who permanently reside in your household.	□ [1] = 1 child □ [2] = 2 children □ [3] = 3 children □ [4] = 4 or more children
21	**How old is your child** that currently goes to [name of participating primary school]?	□ [1] = 5 years or younger □ [2] = 6 years □ [3] = 7 years □ [4] = 8 years □ [5] = 9 years □ [6] = 10 years □ [7] = 11 years or older
22	**What family relationship** do you have with your child that currently goes to [name of participating primary school]?	□ [A] = Mother □ [B] = Father □ [C] = Other: [open-entry]

23	**What age group** would you assign yourself to?	□ [1] = 24 years or younger □ [2] = Between 25 and 34 years □ [3] = Between 35 and 44 years □ [4] = Between 45 and 54 years □ [5] = Between 55 and 64 years □ [6] = Between 65 or older
24	What educational group would you assign yourself to? **Comment:** This question refers to the highest educational degree you have obtained so far.	□ [1] = Apprenticeship or traineeship in the [German] dual system □ [2] = Degree from a [German technical] college □ [3] = Degree from a [German] university of applied sciences and arts □ [4] = University degree □ [5] = No educational degree □ [6] = Other: [open-entry]
	Page break	
25	Could you briefly describe **to what extent your child uses digital technologies & media** on a daily basis? **Comment:** Notes are sufficient. **Examples:** • What devices does your child normally use and how often (e.g. smartphones, tablets, laptops, Smart TVs and so on)? • Does your child own his or her own device or are the devices shared with others? • To what extent is your child supervised while using digital technologies & media?	[Open-entry]

	• What are the main functions and applications that your child normally uses (e.g. learning games, WhatsApp or similar applications, listening to music or audiobooks, watching movies or shows, taking pictures and so on)? • Do you restrict certain devices to age-appropriate usage modes (e.g. content filters, time restrictions and so on)?	

Page break

26.1 - 26.5	In the following you may find a list of **statements** again. For each statement, please indicate **how well this statement applies to you personally or rather to what extent you agree with it**.	□ [-2] = Totally disagree □ [-1] = Rather Disagree □ [0] = Neutral □ [1] = Rather Agree □ [2] = Totally agree
	I think it is positive if primary school children already learn how they can use digital technologies & media! [Positively poled item]	
	It is quite alarming that nowadays primary school children grow up with so much technology! I think less would be better! [Negatively poled item]	
	I think that nowadays primary school children should already learn basic programming skills! [Positively poled item]	
	In my opinion the most important thing is that primary school children should learn the implicit dangers and risks of using digital technologies & media! [Negatively poled item]	
	Access to digital technologies & media is important so children do not lose touch with their social environment! [Positively poled item]	

Page break

27	[If item 10 = A] You have mentioned earlier that you or other family members use digital voice assistants. Could you briefly describe **to what extent your child uses this voice assistant** on a daily basis? **Comment:** Notes are sufficient. **Examples:** • What are typical situations in which your child uses the digital voice assistant? • What are the main functions and applications that your child normally uses (e.g. music or audiobooks, applications, interactive games and so on)? • Who is normally present in these situations (e.g. family members, friends, child alone)? • Are there certain things that you have noticed when your child uses the voice assistant (e.g. speech, behaviour, learning effects, excitement etc.)?	[Open-entry]

	Page break	
28	[If item 10 = B or C] and [if item 16 = A] and [if item 17 = A] You have mentioned earlier that **your child** has the opportunity **to engage with digital voice assistants** in your **closer circle of family and friends**. Could you **briefly describe** how these situations normally look like? **Comment:** Notes are sufficient. **Examples:** • Can your child actively use the voice assistant or is he or she only passively observing? • What are the main functions and applications that your child normally uses (e.g. music or audiobooks, applications, interactive games and so on)? • Who is normally present in these situations (e.g. family members, friends, child alone)? • Are there certain things that you have noticed when your child uses the voice assistant (e.g. speech, behaviour, learning effects, excitement etc.)?	[Open-entry]
	Page break	
29	[Closing section] **Almost done!** [emojis: 💪🎊🎉] You have **successfully** completed the survey! **Thank you** for your participation! Lastly, we would like to ask you whether you are interested that your child participates in **two interactive workshops**, which will take place on **[workshop dates]** in the premises of the [name of participating primary school]. **All relevant details** can be found on the backside of the **survey invitation**.	[None]
30	Are you interested that your child participates in both of these workshops? **Comment:** Please note that your answer is only supposed to help us with the preparation of the workshop, especially regarding the expected number of participating children.	▫ [A] = Yes, I would be interested that my child participates in both of these workshops. ▫ [B] = No, I am not interested that my child participates in these workshops.
31	[If item 30 = A]	Your personally selected code (e.g. ABF4, S5RD and so on): [Open-entry]

	Participation in the workshop requires that your child has to **submit the signed consent form** in the **beginning of the first workshop.** The consent form is attached to the survey invitation. In order to link your anonymous survey data with the workshop data, we would kindly ask you to do the following: • In the following field you can enter a **random four-digit code** based on **numbers as well as letters** (e.g. ABF4, S5RD and so on). • After you have entered a random four-digit code, **please use the same code** on the respective field on the consent form. • In order to ensure the anonymity of your personal data the consent forms will be **treated as confidential** and **destroyed upon the completion of the study**.	
32	**Thank you** for your participation! In case you have **any further questions** please do not hesitate to reach out to us! You may find the **contact details** on the the survey invitation! With best regards [Name student investigator] & [Name principal investigator]	[None]

Notes. Table shows detailed structure of the online questionnaire as it was used in parental survey, including section and page structure, section headlines, item numbers, item contents, item styles (e.g. bold-printed words, emojis) conditionalities and the data entry. Data entries based on letters indicate nominal measurements without implicit ordinal ranks, while data entries based on numbers indicate at least ordinal or even interval measurements. For items 4.1 – 4.16, the original wording used by Karrer et al. (2009) was adjusted, because the key terminology used in their original instrument – namely 'electronical devices' – appeared to be rather outdated and old-fashioned in the context of digitisation in 2019. Hence, the SI made the unorthodox decision adjust the instrument by replacing the term 'electronical devices' with 'digital technologies and media' in the entire instrument. The term 'digital technologies and media' was exemplified in the introduction of the questionnaire (see item 2).

Appendix V. CUREC documents and approval confirmation

Doc No.	Document Copy
1.1	

CENTRAL UNIVERSITY RESEARCH ETHICS COMMITTEE (CUREC)

Form CUREC 1A Checklist for the Social Sciences and Humanities

The University of Oxford places a high value on the knowledge, expertise, and integrity of its members and their ability to conduct research to high standards of scholarship and ethics. The research ethics clearance procedures have been established to ensure that the University is meeting its obligations as a responsible institution.

They start from the presumption that all members of the University will take their responsibilities and obligations seriously and will ensure that their research involving human participants is conducted according to the established principles and good practice in their fields and in accordance, where appropriate, with legal requirements. Since the requirements of research ethics review will vary from field to field and from project to project, the University accepts that different guidelines and procedures will be appropriate.

- Please check **"Where and how to apply for ethical review"** and the **CUREC flowchart** first to see if you need ethics approval.
- Please complete this form using a word processor and email it, together with your supporting documents, to your **Departmental Research Ethics Committee (DREC)** (**if applicable**). If you don't have a DREC please email this form to **ethics@socsci.ox.ac.uk** using your official **ox.ac.uk email** address. **Only emailed applications will be accepted.**

WHAT THIS CHECKLIST IS DESIGNED FOR

This **CUREC 1A checklist** is designed largely for research that falls within the Divisions of Social Sciences and Humanities where ethical issues are relatively few and straightforward. Interviews, field work and oral history are also included in the CUREC process.

The **full CUREC 2 application** is only required where certain project characteristics (e.g. type of participants, or procedures) result in a more complex set of ethical issues. It is expected that only in a limited number of cases will it be necessary for researchers to complete a CUREC 2 application. The checklist below will direct you to a CUREC 2 application if needed.

WHAT THIS CHECKLIST WILL NOT ASSESS

This checklist does not cover research governance, satisfactory methodology, or compliance with the requirements of publishers when administering their tests or questionnaires. As **principal researcher (i.e. principal investigator)**, it is your responsibility to ensure that requirements in these areas are met.

CUREC does not review studies classed as **audit** (see Glossary and Decision Flowchart for CUREC on our website).

If your study involves **NHS patients, NHS staff / data / facilities, or human tissue**, please check the Decision Flowchart for NHS approval and contact the Clinical Trials and Research Governance (CTRG) team in the first instance.

Further information on the University's research ethics procedures is available from the CUREC website.

CUREC 1A Checklist Version 3.1 Approved by CUREC 22 Feb 2018 Page **1** of **12**

1.2

SECTION A: Filter for CUREC2 application This section determines whether your study raises more complex issues which require the completion of a full application for ethical review, known as the CUREC 2 application. **(Please mark 'X' in the Yes/No column as appropriate to indicate your response.)**		
1. Are research participants classed as **people whose ability to give free and informed consent is in question**? (This may include those under 18 (though see "competent youths"), prisoners, or adults "at risk".) Your attention is drawn to the University's **Safeguarding Code of Practice** and its implications for researchers involving children or adults at risk, including the need for the work to be risk assessed and for researchers to undertake related training. (**Note**: If any of your participants are aged 16 or under, please answer 'Yes' here **and** also answer question 5 below.)	Yes	
2. By taking part in the research, will participants be at serious risk of criminal prosecution (e.g. by providing information on drug abuse or child abuse)?		No
3. Does the research involve the **deception** of participants?		No
4. Does your research raise issues relevant to the **Counter-Terrorism and Security Act (the Prevent duty)**, which seeks to prevent people from being drawn into terrorism? Please see advice on this on our **Best Practice Guidance web page**.		No
If you have answered 'No' to all of the questions above please go to **Section B**. If you have answered 'Yes' to any question above continue to question 5 below.		
5. Is your project covered by a CUREC **approved procedure** (formerly known as "CUREC Protocols")?	Yes	
If yes, please give research procedure number(s): **AP25**		
If you answered '**Yes'** to **ANY** of questions **1-4**, **and** answered '**No'** to question **5**, **please stop completing this checklist and do not submit it for ethical review**. Instead, please complete the CUREC 2 application form from the CUREC website. Then submit the CUREC 2 form for ethical review. If you answered 'Yes' to ANY of questions 1-3, **and** answered 'Yes' to question 5, please go on to **Section B**.		

SECTION B: Contact details and project description (NB: must be typed not handwritten)	
Contact details:	
1. **Principal investigator /** supervisor (if student research (title and full name):	Professor Iram Siraj Department of Education University of Oxford 15 Norham Gardens Oxford, OX2 6PY
2. Name of student (if student research):	Janik Festerling St. Antony's College 62 Woodstock Road Oxford, OX2 6JF
3. Degree programme, e.g. DPhil, BA, MPhil, BSc, MSc (if student research):	M.Sc. Education (Research Design & Methodology)
4. Department or Institute name:	Department of Education
5. Address for correspondence (if different from above):	See above

1.2

6. University e-mail (**not** private email) and telephone:	Email: [redacted] Phone: [redacted]
7. Name and status of others taking part in the project, e.g. third year undergraduate; postdoctoral research assistant:	None

SECTION B continued	
Project description:	
8. Title of research project:	Changing Nature of Childhood Environments: Children's Ontologies About Humanoid Artificial Intelligence
9. List of location(s) where project will be conducted:	German Elementary School
10. If your research involves overseas travel or fieldwork and your department requires a travel risk assessment, will you have completed and returned a risk assessment form beforehand? (This has to be approved by your department before you travel. If you are travelling overseas, you are strongly advised to take out University travel insurance.)	Yes
11. Anticipated duration of research project overall:	1 month
12. Anticipated start and end dates of the research project involving human participants:	From: 23/04/19 To: 30/09/19 Please note that you will need ethics approval **before** you start your research. CUREC 1As may take up to 30 days to process.
13. External organisation funding the research (if applicable): None	
14. Title and very brief and simple lay description of **research** (about 150 words), plus description (about 200 words) of the nature of participants.	
a) Title, brief description of research (150 words) in lay language. When describing the research, please include your methodology, how you are applying professional guidelines, and the use to which results/data will be put. **Please also declare any conflicts of interest here.**	

1.3

Title: *Changing Nature of Childhood Environments: Exploring Children's Ontologies about Digital Voice Assistants*

Although there is an increasing presence of digital voice assistants (DVAs) in the lives of young children (e.g. *Siri, Alexa, Google, Cortana* etc.), hardly any scientific studies have been carried out so far regarding their influence on children's development. Therefore, this research aims to gain a first understanding of children make sense of DVAs and their humanoid artificial intelligence. In order to investigate this, the study follows a mixed methods approach (mainly survey and focus groups). The data collection will take place within one or two German primary school settings. Data collection will be conducted by myself and is divided into two consecutive phases:

- Phase 1: Familiarization with school context and written invitation for parents to take part in an anonymous online survey that comprises questions regarding the household's "tech-environment", parental "tech-attitude", and children's "tech-engagement".
- Phase 2: Two focus groups at the school with a group of children (ca. 15 to 20 per school across grades 1 to 4) and one or two volunteering teachers for assistance and safeguarding purposes.

After the data collection, audio-visual recordings from the focus groups will be transcribed and used for a thematic analysis. For the sake of triangulation findings will be complemented with results from the parental online survey and observational data.

SECTION B continued

b) Description of participants and **obtaining informed consent** (200 words). When describing participants, please include

- criteria for inclusion/exclusion
- method of recruitment
- processes for consent to participate

Please ensure you attach as separate documents (if applicable, in English translation):

- your recruitment and advertisement material e.g. a poster or brief invitation letter/ email
- information for participants to read (or hear) before they agree to take part e.g. written information sheets or (only if applicable) oral information scripts.
- a document to record informed consent. Templates for written consent forms and/or oral information scripts (in case of an oral consent process) are available from the CUREC website
- a guide to interview questions (this may be a list of questions to be asked, or a preliminary scope of questions), or a sample of other instruments (such as a sample questionnaire)
- (if relevant) debriefing document after participants have taken part

Recruitment of schools:

Based on personal and pre-existing network relations, the headteachers of three elementary schools in Germany will be contacted. Then, based on individual head teacher consent the schools will be chosen for the study. If all head teachers consent, one or two schools will be selected based on publicly available demographic criteria regarding their presumed student and parent populations.

During the initial familiarization phase, I will be present in the school for one or two weeks in order to get to know the context and the setting. Based on individual verbal teacher consent, I will passively observe classes across grades and subjects. Only written field notes will be taken during this initial phase.

1.4

Recruitment of parents:

After the initial familiarization phase, all students in the school will receive a written invitation for parents/guardians to take part in a voluntary anonymous online survey (survey items are outlined in the appendix to this form). The written invitation will inform parents about the purpose of the study with the use of simple German language.

Recruitment of students:

This part of the study involves normally developing elementary school students of (i.e. across grade 1 to 4 and between the age of 6 to 11). Their recruitment is based on an opt-in procedure: in the written survey invitation for parents/guardians as well as at the end of the online survey itself, parents are asked whether they want their child/their children to participate in two focus groups which take place on two different days after school with the presence of at least one teacher from the same school. If they want their child/their children to participate, they are asked to fill out the consent form that came with the written survey invitation.

Before the focus groups, individual verbal consent is sought from each child before participation. The focus groups will last for about 90 minutes and are audio-visually recorded. One or two volunteering teachers from the school(s) will be present as moderators/assistants.

Criteria for inclusion/ exclusion of students based on parental consent:

Due to general resource restrictions of an M.Sc. project, the focus group will have a maximum size of 20 students. During the focus group, participants will be devided into three smaller sub-groups, and each one will be guided by a moderators/assistants (i.e. researcher or teacher).

If more than 20 parents/guardians consent that their child/their children may participate in the focus group, students will be selected based on information from parental survey in order to diversify the group structure. Parents will be informed if their child – despite their consent – cannot participate in the focus groups and they will be reassured that the reason behind the exclusion is due to external determinants and not of discriminatory nature.

No material or monetary rewards will be given to parents/guardians, children or teachers. A brief written report about the research findings will be provided to the school(s). In addition, on completion of the study I will offer to present any results in an accessible manner to the school and/or parents.

All information and consent sheets are based on the AP-25 templates.

15. What are the ethical issues connected with your research and what steps have you taken to address them? Please do not answer 'none'. The committee needs to see evidence that you have identified potential ethical issues with respect to your research and have taken steps to address them. These issues could relate to:

- your own physical and psychological safety as a researcher (please see the University's and Social Science Division's Safety in Fieldwork guidance
- participant burdens and/or risks, and
- data protection/ confidentiality (please also see section 18).

For more guidance on ethical issues, please see
http://researchsupport.admin.ox.ac.uk/governance/ethics/resources

1.5

Safeguarding during focus groups:

For the participation in the focus groups, children will remain on school premises after the end of their school day. Therefore, sufficient safeguarding must be provided before the start of each focus group (e.g. while children are arriving at the venue), during the focus groups (e.g. safeguarding for children who do not give individual verbal consent despite their parents written consent) and after the focus group (e.g. making sure that all children are picked up by parents or on their usual way back home). In order to address these safeguarding issues, one or two volunteering teachers from the school will assist during this entire focus group process.

Informed consent:

Parents might feel "obliged" to consent to their child's/their children's participation due to the fact that the focus groups will take place in a school context and that one or two teachers from the same school will be present. Therefore, parental consent forms will make clear that that participation is a completely voluntary extra curriculuar activity and without any academic rewards.

Interactions within children during the focus groups:

During the activities of the second focus group, children might refer to personal experiences regarding their engagement with modern technologies at home (e.g. their engagement with *Alexa* or *Google*). Among children themselves, this could trigger jealousy or feelings of being deprived due to unequal access to modern technologies at home. Therefore, during the moderation of the focus group special attention will be paid to avoid that conversations become too explicit with respect to unequal experiences of technological engagement at home.

Anonymity and confidentiality

See question 18 for details.

Section B continued		
16. Will you obtain informed consent according to CUREC guidelines and good practice in your discipline before participation?	Yes	
17. Will your research involve discussing sensitive issues? This could be information relating to race or ethnic origin, political opinions, religious beliefs, physical/mental health, trade union membership, sexual life or criminal activities.		No
18. **Management and handling of personal and other research data** Your management and handling of personal data and special category data of human participants, either directly or via a third party, will need to comply with the requirements of the General Data Protection Regulation (GDPR) and the new Data Protection Act, as set out in the University's Guidance on Data Protection and Research. In answering the questions below, please also consider the points raised in the Data Protection Checklist. For advice on research data management and security, please consult with the University's Research Data Team (**researchdata@ox.ac.uk)** and/or your local IT department and the University's web pages on research data management.		
a) Will your research involve the collection of **records of consent** (e.g. written forms, audio-recorded, or other recorded consent)? **If 'Yes', these will be classed as fully identifiable personal data (directly linked to an individual).**	Yes	

1.6

b) Will your research involve the collection of **other** **personal data?** **If 'Yes**', specify in what form(s) this will be stored: • Fully identifiable (directly linked to an individual) • Pseudonymised (potentially identifiable as data may be attributed to an individual if linkage information can be accessed elsewhere by researchers) • Fully anonymised (i.e. cannot be linked to an individual)	Yes Yes Yes Yes	
c) Will any of the personal data you collect classify as **special category** data? **If 'Yes',** in what form(s) will this be stored: • Fully identifiable (directly linked to an individual) • Pseudonymised (potentially identifiable as data may be attributed to an individual if linkage information can be accessed elsewhere by researchers) • Fully anonymised (i.e. cannot be linked to an individual)		No
d) How will any personally identifiable data be collected, transferred and backed up? Please describe the arrangements for any physical transfer of personal data (including paper records and data captured electronically via portable media) from where it is collected to local storage.		
The online survey will be conducted with a *Qualtrix* questionnaire. While the survey is open for respondents, all *Qualtrix* data are kept on servers in the EU. Once the official time period for the online survey is over, I will download the data from the *Qualtrix* website and store it on an external hard drive with password protection, so it is no longer on remote servers. If parents indicate that they want their child/their children to participate in the two focus groups, they are asked to put down the child's first name and the first initial of the surname in a text field in the questionnaire (e.g. Janik Festerling > Janik F.). After downloading the survey data, I will anonymise this data using pseudonyms and code numbers. The identifiers will be stored separately in a locked excel file (see f). Focus groups will be audio-visually recorded and transcribed in the aftermath of the focus groups. In the transcripts the pseudonyms will be used to link information to the online survey data. Recordings, transcripts and all other data will be stored on an external hard drive with password protection. For the entire data analysis and the reporting of results only pseudonymised data will be used in order to ensure anonymity, as is suggested by the protocol for Approved Proceedure number 25 and by the BERA (2018) ethical guidelines (p. 21). No corporate or georgaphical identifiers will be used that may compromise anonymity of the students, the teachers and the schools.		
e) Where, and for how long, will participants' personally identifiable data be stored during and after the study? (Please outline the procedures for ensuring confidentiality, e.g. security arrangements, anonymisation or pseudonymisation of such data. Please distinguish between records of consent and other forms of personally identifiable data stored)		

1.7

Personally identifiable paper records (including signed consent forms) and the excel file will be kept in a locked cabinet in my college accommodation and destroyed after the study has been completed. Fully pseudonymised transcribed data from the audio-visual recordings and questionnaire data will be kept on my computer in a password protected encrypted file and backed up in a password protected encrypted file on an external hard drive. Only the fully pseudonymised data will be retained after the project for further research, this is deemed acceptable by AP-25 which states that "There is no time limit on retention of completely anonymised data. If non-anonymised data is to be retained, the consent form should seek consent for this retention" (2018, p. 9).
f) If storing pseudonymised data, please confirm that identifiers will be held separately from the research data and linked through a unique study number. Specify how and at what point the pseudonymisation will occur, how the linkage information will be stored and state whether or not (and when) the linkage will be destroyed.
A separate excel file that comprises the linkage between pseudonyms and the children's first name and the surname's initial will be kept on its own encrypted USB drive (device details: 16 GB iStorage datAshur with AES 256-bit hardware encryption).
g) Who will have access to the personally identifiable data? If personally identifiable data is to be shared with another organisation, how will it be transferred/disclosed securely?
I will be the only person who has access to the data file mentioned under (f) as well as the audio-visual recordings. This will be explicitly stated in information sheets and consent forms.
h) When and how will personally identifiable data be destroyed? (NB. Personally identifiable data should be destroyed when no longer required.)
The data file mentioned under (f) as well as audio-visual recordings will be destroyed upon completion of my M.Sc.
i) How, where and for how long will other research data be stored after the study has finished? For more information about University and research funder retention policies, please see the University's web pages on research data management.
For potential future research, the pseudonymised data will be stored on a hard drive in a password protected file. If the research is published pseudonymised data and records will be stored for a minimum retention period of 3 years after publication or public release of the work of the research.

1.8

SECTION C: Methods and procedures to be used	
Method used: Please ensure you have addressed any potential ethical issues related to these methods in Section 14 and in your Participant Information Sheet	**Please mark 'X'**
1. Analysis of existing records	
2. Snowball sampling (recruiting through contacts of existing participants)	
3. Use of casual or local workers e.g. interpreters	
4. Participant observation	Yes
5. Covert observation	
6. Observation of specific organisational practices	
7. Participant completes questionnaire in hard copy	
8. Participant completes online questionnaire or other online task	Yes
9. Using social media	
10. Participant performs paper and pencil task	
11. Participant performs verbal or aural task (e.g. for linguistic study)	
12. Focus group	Yes
13. Interview	
14. Audio recording of participant (you will generally need specific consent from participants for this)	Yes
15. Video recording of participant (you will generally need specific consent from participants for this)	Yes
16. Photography of participant (you will generally need specific consent from participants for this)	
17. Others (please specify):	

1.9

SECTION D: Professional guidelines and training		
In this section, please mark 'X' against at least one of the following professional guidelines you aim to adhere to. You should use the principles listed in your chosen guideline(s) in conducting your own research. **Note**: this is not an exhaustive list.		**Please mark 'X'**
Research specialism/ methodology	**Association and guidance document**	
Anthropology	Association of Social Anthropologists of the UK and Commonwealth	
Criminology	http://www.britsoccrim.org/ethics/	
Education	British Educational Research Association Ethical Guidelines for Educational Research	X
Geography	Association of American Geographers Statement on Professional Ethics	
History	Oral History Society of the UK Ethical Guidelines	
Internet-based Research	British Psychological Society: Conducting Research on the Internet Association of Internet Researchers Ethics Guide Also see our Best Practice Guidance on internet-based research	
Law (Socio-Legal)	Socio-Legal Studies Association: Statement of Principles of Ethical Research	
Management	Academy of Management's Professional Code of Ethics	
Political Science	American Political Science Association (APSA) Guide to Professional Ethics in Political Science	
Politics	Political Studies Association. Guidelines for Good Professional Conduct	
Psychology	British Psychological Society Code of Ethics and Conduct	
Social Research	Social Research Association: Ethical Guidelines	
Sociology	The British Sociological Association: Statement of Ethical Practice	
Visual Research	ESRC National Centre for Research Methods Review Paper: Visual Ethics: Ethical Issues in Visual Research	
Other professional guidelines. Please specify the other guidelines used here:		

Please indicate what training in research ethics the researchers involved with this study have received, e.g. the title of the course and date completed (online training available at http://researchsupport.admin.ox.ac.uk/support/training/ethics). If no formal training has been undertaken, please indicate any discussions of research methodology between researchers and supervisors here.
Lecture on research ethics during Michaelmas 2018 (Course: Foundations of Educational Research I) Two summative essays on research ethics (Title & mark: "The *True* Challenge of Ethically Responsible Research in the 21st Century", *distinction*; "A Treatise on the Hum(e)an Nature of the Researcher-Researched Relationship", *high distinction*)

1.10

SECTION E: Signatures (The SSH IDREC Secretariat accepts either option below. If you have a DREC, check which signature option it prefers.)

- **Option 1:** 'Electronic signatures', i.e. email confirmations from a University of Oxford email address, can be accepted. Separate emails should come from each of the relevant signatories as outlined below, indicating acceptance of the relevant responsibilities. **Pasted images of signatures cannot be accepted in the sections below.**
- **Option 2:** Handwritten (wet-ink) signatures. Please scan them and the rest of the checklist pages to create a single PDF document and email through.

Please ensure this checklist is signed by:

For staff research:	For student research:
1. Principal investigator	1. Principal investigator (project supervisor)
2. Head of Department (or nominee)	2. Head of Department (or nominee)
	3. Student researcher

1. Principal investigator signature/supervisor signature (if student research)

I understand my responsibilities as **principal investigator** as outlined in the CUREC glossary and guidance on the CUREC website.

I declare that the answers above accurately describe the research as presently designed, and that a new checklist will be submitted should the research design change in a way which would alter any of the above responses so as to require completion of CUREC 2 (involving full scrutiny by an IDREC). I will inform the relevant IDREC if I cease to be the principal investigator on this project and supply the name and contact details of my successor if appropriate.

Signature: .....

Print name (block capitals): ..... **Date**: .....

2. Departmental endorsement signature

I have read the research project application named above. On the basis of the information available to me, I:

(i) consider the principal investigator to be aware of her/his ethical responsibilities in regard to this research;
(ii) consider that any ethical issues raised have been satisfactorily resolved or are covered by relevant professional guidelines and/or CUREC approved procedures, and that it is appropriate for the research to proceed (noting the principal investigator's obligation to report should the design of the research change in a way which would alter any of the above responses so as to require completion of a CUREC 2 full application);
(iii) am satisfied that: the proposed project design and scientific methodology is sound; the project has been/will be subject to appropriate **peer review**; and is likely to contribute to existing knowledge and/or to the education and training of the researcher(s) and that it is in the **public interest**.

Signed by Head of Department or nominee *(example nominees for student research include the Director of Graduate Studies/ Director of Undergraduate Studies)*:

Signature: .....

Print name (block capitals): ..... **Date**: .....

3. Student signature (if student research)

I understand the questions and answers that have been entered above describing the research, and I will ensure that my practice in this research complies with these answers, subject to any modifications made by the principal investigator properly authorised by the CUREC system.

Signed by student: Janik Festerling

Print name (block capitals): JANIK FESTERLING **Date**: 05/03/2019

L

1.11

SECTION F: SUBMITTING THE COMPLETED CHECKLIST	Please mark 'X'
1. Check you have completed all sections (A-E)	Yes
2. Ensure your application is signed by you, your supervisor (if student) and department	Yes
3. **Please attach all supporting documents (see section B, question 14b for details).** If the appropriate supporting documentation is not included with your application, you will then be asked to provide this separately. **This may well delay the ethical review process, and thus the start of your research.**	Yes
4. Ensure you have declared conflicts of interest (if any) in Section B, question 14a.	Yes
5. If your department has a Departmental Research Ethics Committee (DREC), submit this checklist and supporting information to the appropriate departmental officer.	Yes
6. If your department does not have a DREC, submit the checklist and supporting information to the SSH IDREC (email ethics@socsci.ox.ac.uk).	N.A.
7. Applications must be sent by email from your official ox.ac.uk email account. Please do not send applications by post.	Yes

LI

2.1

Subject: CUREC
Date: Tuesday, 26 March 2019 at 09:19:13 Greenwich Mean Time
From: Liam Gearon
To: Janik Festerling
CC: Iram Siraj, Education Research Office

Dear Jank,

(Please note applications should be sent initially to the Education Research office who then distributes the applications to selected members of the Committee)

Changing Nature of Childhood Environments: Children's Ontologies About Humanoid Artificial Intelligence

The above application has been considered on behalf of the Departmental Research Ethics Committee (DREC) in accordance with the procedures laid down by the University for ethical approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to DREC, the proposed research has been judged as meeting appropriate ethical standards, and accordingly, approval has been granted.

If your research involves participants whose ability to give free and informed consent is in question (this includes those under 18 and vulnerable adults), then it is advisable to read the following NSPCC professional reporting requirements for cases of suspected abuse

http://www.nspcc.org.uk/globalassets/documents/information-service/factsheet-child-abuse-reporting-requirements-professionals.pdf

Should there be any subsequent changes to the project which raise ethical issues not covered in the original application you should submit details to research.office@education.ox.ac.uk for consideration.

Good luck with your research study.

Sincerely,

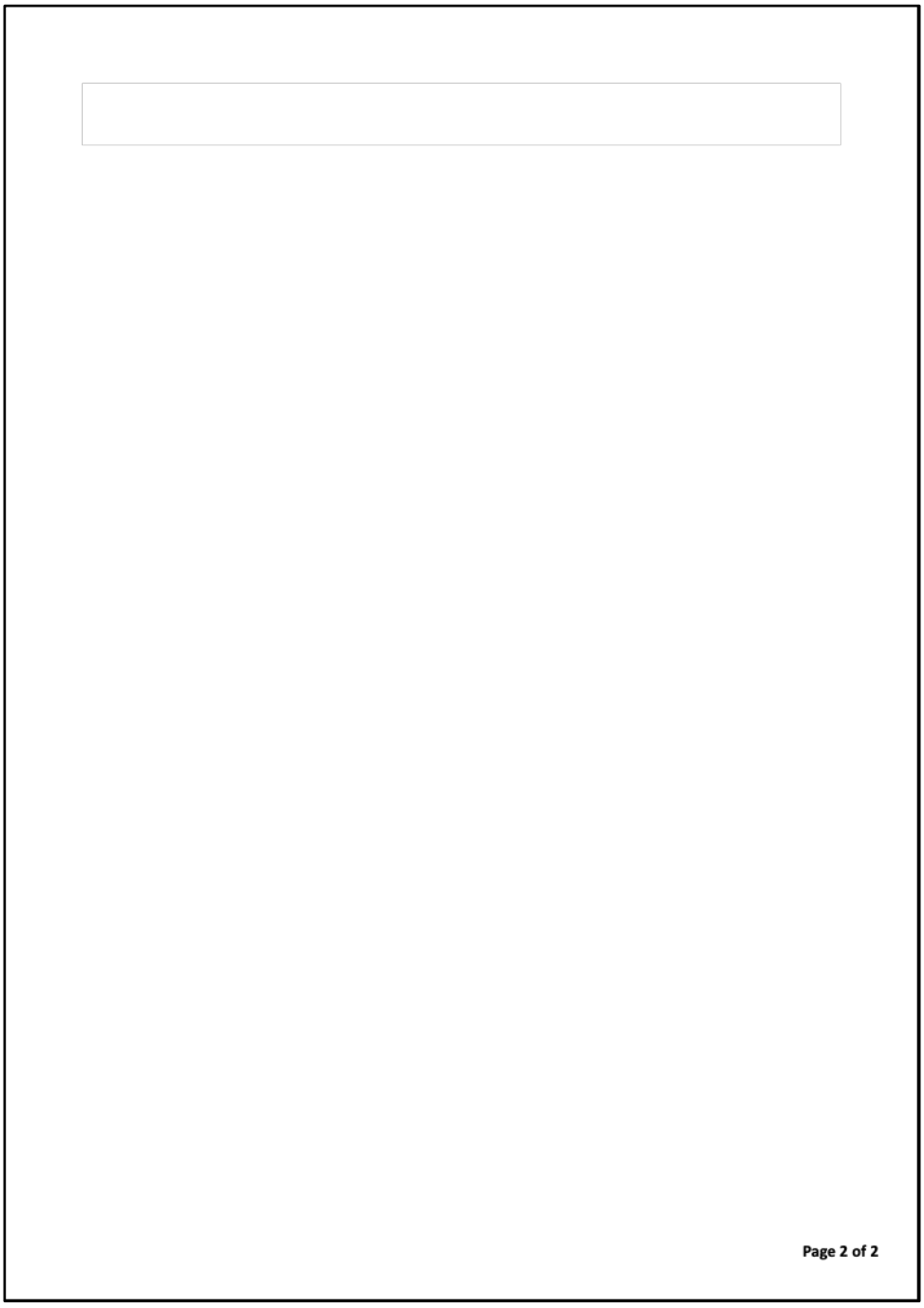
With kind regards,

Liam

Dr LF Gearon

Chair DREC
Department of Education, University of Oxford

Page 1 of 2

2.2	

Notes. First part of the table (1.1 – 1.12) shows electronic copies of CUREC documents that were submitted to the DREC. Second part of the table (2.1 – 2.2) shows electronic copy of the approval, which was granted via email. A handwritten signature was provided by the SI (1.11), and electronical signatures were provided by the PI as well as the chair of the DREC (2.1).

Appendix VI. Frequencies of demographical characteristics in parental survey

Characteristic	Total Sample (*n*=67)	Focus Groups (*n*=26)	School 1 (*n*=36)	School 2 (*n*=31)
Number of children living in the household				
1 child	17 (25.4%)	5 (19.2%)	9 (25.0%)	8 (25.8%)
2 children	36 (53.7%)	16 (61.5%)	18 (50.0%)	18 (58.1%)
3 children	11 (16.4%)	3 (11.5%)	7 (19.4%)	4 (12.9%)
4 or more children	3 (4.5%)	2 (7.7%)	2 (5.6%)	1 (3.2%)
Child's age				
6 years	10 (14.9%)	3 (11.5%)	4 (11.1%)	6 (19.4%)
7 years	10 (14.9%)	1 (3.8%)	6 (16.7%)	4 (12.9%)
8 years	19 (28.4%)	8 (30.8%)	12 (33.3%)	7 (22.6%)
9 years	14 (20.9%)	6 (23.1%)	10 (27.8%)	4 (12.9%)
10 years	14 (20.9%)	8 (30.8%)	4 (11.1%)	10 (32.3%)
Respondent's family relationship with child				
Mother	52 (77.6%)	19 (73.1%)	26 (72.2%)	26 (83.9%)
Father	13 (19.4%)	6 (23.1%)	9 (25.0%)	4 (12.9%)
Other	2 (3.0%)	1 (3.8%)	1 (2.8%)	1 (3.2%)
Respondent's age group				
25 – 34 years	10 (14.9%)	2 (7.7%)	7 (19.4%)	3 (9.7%)
35 – 44 years	42 (62.7%)	16 (61.5%)	24 (66.7%)	18 (58.1%)
45 – 54 years	14 (20.9%)	8 (30.8%)	5 (13.9%)	9 (29.0%)
55 – 64 years	1 (1.5%)	0 (0%)	0 (0%)	1 (3.2%)
Respondent's highest educational degree				
Apprenticeship or traineeship	20 (29.9%)	7 (26.9%)	15 (41.7%)	5 (16.1%)
Technical college	6 (9.0%)	4 (15.4%)	4 (11.1%)	2 (6.5%)
Applied sciences and arts	18 (26.9%)	4 (15.4%)	8 (22.2%)	10 (32.3%)
University	19 (28.4%)	8 (30.8%)	8 (22.2%)	11 (35.5%)
No educational degree	1 (1.5%)	0 (0%)	0 (0.0%)	1 (3.2%)
Other	3 (4.5%)	3 (11.5%)	1 (2.8%)	2 (6.5%)

Notes. Table shows frequencies of demographical characteristics in parental survey, including values for the total sample (first column), for the focus group sample (second column) and for both schools (third and fourth column). Children's demographical characteristics always refer to the child that was enrolled in school 1 or 2 at the time of data collection. Demographical characteristics with frequencies of zero in the total sample are excluded from the table. Both observational units who chose the category 'Other' for 'Respondent's family relationship with child' indicated that both of the parents responded to the survey together. One case was excluded from the analysis due to missing values. Educational categories correspond to the German education system. All percentages are calculated per column and within each demographical characteristic.

Appendix VII. Frequencies of other personal characteristics in parental survey

Characteristic	Total Sample (*n*=68)	Focus Groups (*n*=26)	School 1 (*n*=37)	School 2 (*n*=31)
Prevalence of technological trends and developments at work				
Rarely	5 (7.4%)	2 (7.7%)	1 (2.7%)	4 (12.9%)
Occasionally	13 (19.1%)	5 (19.2%)	9 (24.3%)	4 (12.9%)
Often	28 (41.2%)	10 (38.5%)	11 (29.7%)	17 (54.8%)
Very Often	18 (26.5%)	7 (26.9%)	12 (32.4%)	6 (19.4%)
Currently unemployed	4 (5.9%)	2 (7.7%)	4 (10.8%)	0 (0%)
Familiarity with MI				
Slightly familiar / Have heard of it a once or twice	17 (25.0%)	8 (30.8%)	12 (32.4%)	5 (16.1%)
Somewhat familiar / Hear it from time to time	27 (39.7%)	10 (38.5%)	12 (32.4%)	15 (48.4%)
Moderately familiar / Have read up on it a little bit	20 (29.4%)	8 (30.8%)	10 (27.0%)	10 (32.3%)
Extremely familiar / Know a lot about it	4 (5.9%)	0 (0.0%)	3 (8.1%)	1 (3.2%)

Notes. Table shows frequencies of other personal characteristics in the parental survey, including values for the total sample (first column), for the focus group sample (second column) and for both schools (third and fourth column). Personal characteristics with frequencies of zero in the total sample are excluded from the table. All percentages are calculated per column and within each personal characteristic.

Appendix VIII. Summary statistics of attitudinal survey instruments

	Total Sample (*n*=68)			Focus Groups (*n*=26)			School 1 (*n*=37)			School 2 (*n*=31)		
Survey Instrument	M (SD)	Min	Max	M (SD)	Min	Max	M (SD)	Min	Max	M (SD)	Min	Max
General towards attitude digital technologies and media	.12 (.51)	-1.25	1.25	.02 (.55)	-1.25	1.25	.16 (.55)	-1.25	1.25	.07 (.46)	-.94	1.00
Attitude towards MI	.03 (.75)	-1.40	2.00	-.01 (.78)	-1.40	2.00	-.12 (.71)	-1.40	1.80	.21 (.75)	-1.00	2.00
Attitude towards children's engagement with digital technologies and media	-.01 (.63)	-1.80	1.60	-.04 (.70)	-1.60	1.60	-.01 (.62)	-1.60	1.40	-.01 (.63)	-1.80	1.60

Notes. Table shows descriptive statistics of all three attitudinal survey instruments (M = Mean, SD = Standard Deviation, Min = Minimum value in the data, Max = Maximum value in the data). The dependent variable for each instrument is the attitudinal average score that was computed for each observational unit (i.e. sum of item responses divided by number of items). Negatively poled items were reversed before average attitudinal scores were computed for each observational unit. The statistical analysis of all attitudinal items assumed that underlying 5-step Likert scales are continuous measurements on an interval level, therefore implying equal underlying real-world differences across all scale-steps (cf. Field, 2009).

Appendix IX. Internal reliability of attitudinal survey instruments

Survey Instrument	Cronbach's Alpha (α)	Cronbach's Alpha Based on Standardised Items (*A*)	Number of Items
General attitude towards digital technologies and media	.84	.85	16
Attitude towards MI	.84	.84	5
Attitude towards children's engagement with digital technologies and media	.71	.72	4

Notes. Table shows Cronbach's Alpha values for all three attitudinal instruments used in the parental survey. 'Cronbach's Alpha (α)' is based on the covariances among items, while 'Cronbach's Alpha Based on Standardised Items (*A*)' is based on correlations among items, therefore assuming that all items have equal variances. For the first and second instrument one case was excluded from the analysis due to missing values. For the third instrument dropping item 26.4 (**Appendix II**) strongly improved the instrument's internal reliability according to Cronbach's Alpha ($\Delta\alpha = 0.08$; $\Delta A = 0.12$).

Interpretation of Cronbach's Alpha. Cronbach's Alpha indicates how strong different items on a scale are interrelated, meaning that it examines the variation across items and subjects (Cortina, 1993). Therefore, Cronbach's Alpha can be used to investigate the internal reliability of scales: higher values indicate stronger interrelation (i.e. higher score on one item is associated with higher score on all other items) and thus a stronger internal reliability of the respective scale (Cortina, 1993). Although absolute value cut-offs should be treated with caution (Cortina, 1993), most recommended value ranges for the acceptance of a particular scale centre around $0.7 < \alpha < 0.9$ (cf. Peterson, 1994).

Appendix X. Correlation among survey instruments

Survey Instrument	1.	2.	3.
1. *General attitude towards digital technologies and media*	1		
2. *Attitude towards MI*	.57	1	
3. *Attitude towards children's engagement with digital technologies and media*	.49	.48	1

Notes. Table shows Pearson correlation coefficients among all three attitudinal instruments used in the parental survey.

Interpretation of Pearson's. The Pearson correlation coefficient indicates how strong different continuously measured variables are associated with each other on average and in a linear way, further implying that positive coefficients indicate a positive linear association (i.e. on average, an increase in one variable is associated with an increase of the other variable, while the value of the coefficient itself indicates the magnitude of this linear relationship); negative coefficients indicate the opposite (Field, 2009).

Appendix XI. Frequencies of DVA-ownership and usage in parental survey

(I) Sample overview	DVA-ownership *and* usage (*n*=19)	DVA-ownership *but no* usage (*n*=32)	*No* DVA-ownership *or* usage (*n*=16)
Focus Groups (*n*=26)	8 (30.8%)	15 (57.7%)	3 (11.5%)
School 1 (*n*=36)	10 (%27.8)	17 (47.2%)	9 (25.0%)
School 2 (*n*=31)	9 (29.0%)	15 (48.4%)	7 (22.6%)
(II) Characteristics			
General attitude towards digital technologies and media			
Low attitude	2 (10.5%)	5 (15.6%)	2 (12.5%)
Average attitude	13 (68.4%)	21 (65.6%)	13 (81.3%)
High attitude	4 (21.1%)	6 (18.8%)	1 (6.3%)
Attitude towards MI			
Low attitude	2 (10.5%)	5 (15.6%)	3 (18.8%)
Average attitude	10 (52.6%)	22 (68.8%)	13 (81.3%)
High attitude	7 (36.8%)	5 (15.6%)	0 (0%)
Attitude towards children's engagement with digital technologies and media			
Low attitude	1 (5.3%)	3 (9.4%)	4 (25%)
Average attitude	15 (78.9%)	25 (78.1%)	11 (68.8%)
High attitude	3 (15.8%)	4 (12.5%)	1 (6.3%)
Respondent's highest educational degree			
Apprenticeship or traineeship	7 (36.8%)	6 (18.8%)	7 (43.8%)
Technical College	1 (5.3%)	5 (15.6%)	0 (0%)
Applied Sciences and arts	3 (15.8%)	8 (25%)	7 (43.8%)
University	7 (36.8%)	10 (31.3%)	2 (12.5%)
No educational degree	0 (0%)	1 (3.1%)	0 (0%)
Other	1 (5.3%)	2 (6.3%)	0 (0%)

Notes. Part (I) of the table shows frequencies of DVA-ownership (i.e. owning at least one device that allows the owner to access a DVA) and usage (i.e. not only owning but also using at least one DVA) in parental survey, including values for the focus group sample (first row) and for both schools (second and third row). Part (II) of the table shows frequencies of DVA-ownership and usage across various personal characteristics. Attitudinal categories were generated based on differences between individual attitudinal scores and the overall sample mean of the attitudinal score (i.e. low/high attitude: individual attitudinal score is below/above 1 SD of sample mean; average attitude: individual attitudinal score is within 1 SD of sample mean). One case was excluded from the analysis due to missing values. Percentages in part (I) of the table are calculated per row (i.e. per group), while percentages in part (II) of the table are calculated per column and within each characteristic.

Appendix XII. Mean comparison of attitudinal survey instruments

	Total Sample (*n*=68)			DVA-ownership *and* usage (*n*=19)			DVA-ownership *but no* usage (*n*=32)			*No* DVA-ownership *or* usage (*n*=16)		
Survey Instrument	M (SD)	Min	Max	M (SD)	Min	Max	M (SD)	Min	Max	M (SD)	Min	Max
General attitude towards digital technologies and media	.12 (.51)	-1.25	1.25	.26 (.57)	-.94	1.19	.09 (.54)	-1.25	1.25	.01 (.36)	-.63	.69
Attitude towards MI	.03 (.75)	-1.40	2.00	.23 (.82)	-1.40	1.80	.05 (.77)	-1.40	2.00	-.24 (.53)	-1.00	.60
Attitude towards children's engagement with digital technologies and media	-.01 (.63)	-1.80	1.60	.05 (.55)	-.80	1.00	.05 (.65)	-1.60	1.60	-.19 (.63)	-1.80	.80

Notes. Table shows mean comparison of all three attitudinal survey instruments (M = Mean, SD = Standard Deviation, Min = Minimum value in the data, Max = Maximum value in the data) and across different combinations of DVA-ownership (i.e. owning at least one device that allows the owner to access a DVA) and usage (i.e. not only owning but also using at least one DVA). Dependent variable for each instrument is the attitudinal average score that was computed for each observational unit (i.e. sum of item responses divided by number of items). One case was excluded from the analysis due to missing values.

Appendix XIII. Code book of thematic analysis applied to parental survey

(I) Children's general engagement with digital technologies and media (open entry survey item 25)

Code name	Code description & Examples	References
Communication		
WhatsApp	Parental respondents indicate that child uses WhatsApp to communicate with family members or friends	17
Phone calls	Parental respondents indicate that child uses mobile devices to call family members or friends	3
Device ownership		
Child owns a device	Parental respondents indicate that child 'owns' an unspecified device (i.e. child does not have to share the device with somebody else)	25
Child owns DVA	Parental respondents indicate that child 'owns' a DVA (e.g. Alexa)	4
Child owns mobile device	Parental respondents indicate that child 'owns' a mobile device (e.g. smartphones, tablet)	17
Child owns other offline devices	Parental respondents indicate that child 'owns' a device without internet access (e.g. laptop without internet access, digital camera, TV)	2
Child owns stationary computer	Parental respondents indicate that child 'owns' a stationary computer (e.g. laptop, desktop computer)	2
Child owns video game console	Parental respondents indicate that child 'owns' a gaming console (e.g. PS4, Nintendo Wii)	2
Parental and family owned devices	Parental respondents indicate that child uses shared devices that are 'owned' by other individuals or the family in general (e.g. parental smartphones or tablets, smart TVs, DVAs, laptops)	40
Devices		
Mobile devices	Parental respondents indicate that their child uses smartphones or tablets (e.g. iPhones, iPads)	56
Offline devices	Parental respondents indicate that their child uses offline devices (e.g. MP3-players, iPods, digital cameras, laptops without internet access)	4
Smart TV	Parental respondents indicate that their child uses television devices that allow users to directly access streaming services (e.g. Netflix or Amazon Prime)	7
Stationary computer devices	Parental respondents indicate that their child uses stationary computer devices (e.g. desktop computers, Laptops)	17
Video game consoles	Parental respondents indicate that their child uses gaming consoles (e.g. PS4, Nintendo Wii, Nintendo DS)	9
Supervision		
Content restrictions	Parental respondents indicate that their child cannot access or use certain contents due to on-device controls or in-app controls (e.g. browser restrictions, YouTube Kids, child-friendly search engines)	12
Direct supervision	Parental respondents indicate that they personally observe their child's engagement with digital technologies and media or even join the child	19
Indirect supervision	Parental respondents indicate that they are generally aware of their child's engagement with digital technologies and media, personally observe it sometimes but not always, or talk about it with the child	8
Restricted device access	Parental respondents indicate that their child can only access designated devices (e.g. tablets or smartphone with content restrictions, laptops without internet access)	5
Prevention of internet access	Parental respondents indicate that internet access is omitted on certain devices (e.g. browser is disabled)	6

Time restrictions		
Daily time restriction	Parental respondents indicate that their child can only use certain devices for a specified period of time per day	18
General time restriction (unspecified)	Parental respondents indicate that their child can only use certain devices for a restricted but unspecified period of time	12
Monthly time restrictions	Parental respondents indicate that their child can only use certain devices for a specified period of time per month	2
Other forms of restrictions	Parental respondents indicate that time restrictions for their child's usage of digital technologies and media are exercised through other means (e.g. vouchers, negotiated individually based on parental gut feeling)	4
Weekly time restriction	Parental respondents indicate that their child can only use certain devices for a specified period of time per week	9
Usage		
Games	Parental respondents indicate that their child uses devices to play games, without explicitly mentioning educational purposes	49
Usage for educational purposes	Parental respondents indicate that their child uses digital technologies and media for educational purposes (e.g. educational games, typing exercises, looking up information)	30
Antolin (reading software)	Parental respondents indicate that their child uses 'Antolin' (i.e. reading app developed by the German publisher Westermann)	5
Educational purposes on mobile devices	Parental respondents indicate that their child uses mobile devices for educational purposes	14
Educational purposes on stationary devices	Parental respondents indicate that their child uses stationary devices for educational purposes	8
Information seeking	Parental respondents indicate that their child uses digital technologies and media to search for specific information (e.g. news or facts related to football, homework)	8
Audiobooks	Parental respondents indicate that child uses digital technologies and media to listen to audiobooks	31
Music	Parental respondents indicate that their child uses digital technologies and media to listen to music	41
Photography	Parental respondents indicate that their child uses digital technologies and media for photography (e.g. taking or looking at pictures)	20
Watching Movies, Shows and Videos	Parental respondents indicate that their child uses digital technologies and media to watch movies, shows or (online) videos	46
Netflix	Parental respondents indicate that their child uses Netflix	6
YouTube	Parental respondents indicate that their child uses YouTube or YouTube Kids	23

(II.A) Children's engagement with DVAs at home (open entry survey item 27)

Code name	Code description & examples	References
Excitement		
Asking funny questions	Parental respondents indicate that their child likes to ask funny questions and wait for DVA-responses (e.g. Alexa, how are you?)	7
Probing adults with DVA-knowledge	Parental respondents indicate that their child likes to ask knowledge-based questions in order to confirm whether adults (e.g. parents) know the answer as well	4
Usage when boredom occurs	Parental respondents indicate that DVA-usage mainly occurs when their child is bored	2

Using funny features	Parental respondents indicate that their child likes to use funny features that the DVA has to offer (e.g. Alexa 'Fart Generator')	3
Supervision		
Direct supervision	Parental respondents indicate that they or other family members personally observe their child's engagement with the DVA or even engage in joint activities	9
Indirect supervision	Parental respondents indicate that they are generally aware of their child's engagement with the DVA, personally observe or hear it sometimes but not always, or talk about it with the child	5
Technologically enabled supervision	Parental respondents indicate that they regularly check the DVA's voice protocol or usage history that can be accessed through mobile devices	4
Usage		
Audiobooks	Parental respondents indicate that their child uses the DVA to listen to audiobooks	6
Information seeking	Parental respondents indicate that their child uses the DVA to listen to music	8
Interactive games	Parental respondents indicate that their child uses the DVA to play interactive games (e.g. 'Akinator', quiz games)	5
Music	Parental respondents indicate that their child uses the DVA to listen to music	10
Practical assistance	Parental respondents indicate that their child uses the DVA for practical assistance (e.g. setting alarms, setting ambient sounds before going to bed)	4
Selecting movies or shows	Parental respondents indicate that their child uses the DVA to listen to music	2
Weather forecasts	Parental respondents indicate that their child uses the DVA to ask for local weather forecasts	4
Other parental observations		
Diminishing levels of excitement	Parental respondents report that their child's excitement for the DVA has diminished over time	2

(II.B) Children's engagement with DVAs in social circle of family and friends (open entry survey item 28)

Code name	Code description & examples	References
Engaging Excitement		
Asking funny questions	Parental respondents indicate that their child likes to ask funny questions and wait for the responses when engaging with a DVA in the closer circle of family and friends (e.g. Alexa, how are you?)	5
General excitement for DVA-technology	Parental respondents indicate that their child is generally very excited about DVAs after having exposure in the closer circle of family and friends	8
Jealousy (child also wants a DVA at home)	Parental respondents indicate that their child has already indicated the wish to have a DVA at home after having exposure in the closer circle of family and friends	2
Using funny features	Parental respondents indicate that their child likes to use funny features and when engaging with a DVA in the closer circle of family and friends (e.g. Alexa 'Fart Generator')	4
Supervision		
Direct supervision	Parental respondents indicate that adults are always present when their child engages with a DVA in the closer circle of family and friends	3
Indirect supervision	Parental respondents indicate that sometimes but not always adults are probably present when their child engages with a DVA in the closer circle of family and friends	7

Usage

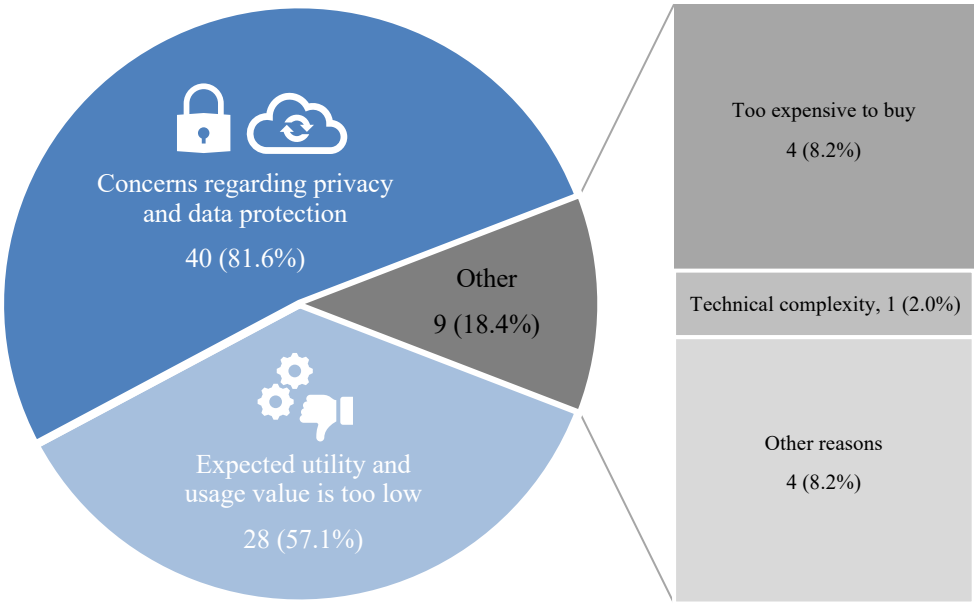
Audiobooks	Parental respondents indicate that their child listens to audiobooks when engaging with a DVA in the closer circle of family and friends	10
Information seeking	Parental respondents indicate that their child seeks for general information from the internet when engaging with a DVA in the closer circle of family and friends	5
Interactive games	Parental respondents indicate that their child plays interactive games when engaging with a DVA in the closer circle of family and friends	4
Music	Parental respondents indicate that their child listens to music when engaging with a DVA in the closer circle of family and friends	2
Practical assistance	Parental respondents indicate that their child uses assisting features when engaging with a DVA in the closer circle of family and friends (e.g. asking DVA to solve an arithmetical problem)	3
Weather forecasts	Parental respondents indicate that their child listens to ask for local weather forecasts when engaging with a DVA in the closer circle of family and friends	3

Other parental observations

Child only observes how others use DVAs	Parental respondents indicate that in the closer circle of family and friends their child passively observes how others engage with DVAs instead of actively engaging with it	10
Diminishing levels of excitement	Parental respondents report that their child's excitement for the DVA has diminished after having DVA-exposure in the closer circle of family and friends for an extended period of time	1

Notes. Tables shows detailed code book of the thematic analysis that was applied to text data of parental survey (see open entry items 25, 27 and 28 in **Appendix II**). Sections I, II.A and II.B of the table are organised by overarching themes that were inductively identified in the data. Each row represents a node within a theme, including the node's name, its description, examples as well as its frequency in the data. The analysis was conducted in NVivo 12.

Appendix XIV. Reasons for refraining from using DVAs among non-users



Notes. Figure visualises the frequency of different reasons for refraining from using DVAs in the group of DVA non-users (n=49). This group includes all parental survey respondents who did not use DVAs despite owning devices that would have allowed them to so, or survey respondents who did not even own such devices. Data entry options with frequencies of zero are excluded from the figure. All percentages are calculated with respect to the sample size of the respective group (sum of percentage values does not equal 100% due to multiple choice entry).

Appendix XV. Code book of thematic analysis applied to focus group data

(I) Children's verbal characterisation of human nature (focus group session 1)

Code name	Code description & Examples	References
Characteristics of human nature		
Human as a sensual being	Interviewee mentions characteristics related to or involving the human senses (i.e. sight, hearing, smell, taste and touch)	7
Human as an empathetic being	Interviewee mentions characteristics related to emotions that humans can feel for one another	2
Human as an emotional being	Interviewee mentions characteristics related to emotions and feelings (e.g. anger, disgust, fear, happiness, sadness, or surprise)	8
Human speech and language	Interviewee mentions characteristics related to human language (e.g. special characteristics of human language)	16
Differences among humans		
Differences in preferences	Interviewee mentions different tastes that humans have (e.g. favourite colours or music)	4
Different body sizes	Interviewee mentions that humans have different body sizes (e.g. some humans are taller, and some humans are shorter)	5
Different eye colours	Interviewee mentions that humans have different eye colours	1
Different faces	Interviewee mentions that each human being has a unique face and facial expression	3
Different voices	Interviewee mentions that human being has a unique voice	2
Gender	Interviewee mentions that differences related to genders (e.g. boys and girls)	3
Human body		
Body parts	Interviewee mentions characteristics related to different parts of the human body (e.g. legs, arms, head, hair)	3
Organs	Interviewee mentions characteristics related human organs (e.g. heart, skin, brain)	5
Human category within the ontological belief system		
Human as a living entity	Interviewee describes human as an animate being	4
Breath	Interviewee describes breathing as an indicator of human animacy	6
Human as a mammal	Interviewee describes that humans belong to the animate group of mammals	7
Human life as a finite existence	Interviewee describes that human life is limited to a certain period of time	2
Motion	Interviewee describes motion of the body as an indicator of human animacy	14
Human as part of nature's ecosystem		
Human as the result of evolution	Interviewee refers to the theory of evolution (e.g. humans descended from apes)	8
Human-animal comparison	Interviewee compares and points out differences and similarities between humans and other animals (e.g. different use of language)	15
Plants	Interviewee refers to plants as another category of living entities or that we need plants in order to survive on the planet (i.e. photosynthesis)	4

Simple organisms	Interviewee refers to simple organisms as another category of living entity (e.g. bacteria)	3

(II) Children's verbal characterisation of HI and MI (focus group session 1)

Code name	Code description & Examples	References
Dimensions of HI		
Smartness	Interviewee characterises intelligence in terms of being smart, including characteristics that indicate how smart somebody is	7
Mathematical thinking	Interviewee mentions characteristics of HI related to mathematical thinking (e.g. mental arithmetic, counting)	5
Memory and knowledge	Interviewee mentions characteristics of HI related to the memorisation of information and the acquisition of knowledge	8
Task-orientation	Interviewee mentions characteristics of HI related to solving tasks and providing answers to questions	2
Nature of learning		
Learning through listening	Interviewee characterises learning as a process of listening to something or someone	3
Learning through repetition and practice	Interviewee characterises learning as process of repeating and practicing exercises	4
Learning through understanding	Interviewee characterises learning as process that requires the learner to understand something	2
Nature of the human brain		
Human brain as a control unit	Interviewee characterises the human brain as control unit (e.g. controlling bodily functions or motions)	11
Human brain as the origin of thought	Interviewee states that humans need their brain to think, belief or reason	6
Perception of machines		
Artefactual language	Interviewee's choice of words while describing machines (e.g. computers) refers to artefactual and object- qualities (e.g. storage, turning machines on or off, hard drives, hard disks, computer chips, programming)	9
Machines as memory systems	Interviewee refers to machines (e.g. computers) as entities that can store information and allow humans to access these information	9
Machines as programmable entities	Interviewee refers to machines (e.g. computers) as entities that are instructed by humans who write programs	10
Perception of the HI-MI relation		
Human brain as a machine	Interviewee uses analogies between the human brain and machines (e.g. computers) regarding their working mechanisms or functions (e.g. human brain is something like a storage disk)	4
Firm perception of ontological HI-MI hierarchy	Interviewee argues that MI can never be more intelligent or more capable than the HI who built it	19
Flexible perception of ontological HI-MI hierarchy	Interviewee is not sure or even opposes that MI can never be more intelligent or more capable than the HI who built it	9

(III) Children's behavioural and verbal engagement patterns (focus group session 2)

Code name	Code description & Examples	References
Volatile dynamics		

Interfering voice commands	Interviewees 'steal' each other's wake words by quickly adding voice command after someone else said the wake word	9
Rapidly changing content	Interviewees quickly change content played by DVA (e.g. changing songs after a couple seconds of playtime)	16
Engaging excitement		
Amusement based on misunderstandings with DVAs	Interviewee displays amusement (e.g. giggling or laughing) based on a misunderstanding with a DVA (e.g. DVA-response to not correspond with intended voice command)	5
Playfulness	Interviewee displays playful engagement (e.g. giggling, laughing, amused or exciting tones)	4
Provoking obscene responses from DVAs	Interviewee uses voice command that is supposed to evoke vulgar or indecent DVA-responses (e.g. asking the DVA to make farting or burping noises)	16
Tone of engagement patterns		
Complementing	Interviewee pays DVA a compliment	4
Courtesy and politeness	Interviewee uses language cues related to courtesy and politeness (e.g. 'thank you' or 'please')	8
Requesting silence	Interviewee tells DVA to be quiet via a voice command	10
Unfriendly or aggressive behaviour	Interviewee displays explicit offensive behaviour towards DVA (e.g. insults or showing a middle finger)	22
Usages and applications		
Asking for explicit content	Interviewee asks DVA to play adult content (e.g. obscene rap songs)	5
Asking for interactive features	Interviewee requests DVA to start interactive features (e.g. 'Alexa, tell me a joke!' or 'Alexa, can you play games?')	13
Asking for music or audiobook content	Interviewee requests DVA to play music songs or audiobooks	17
Asking for the time	Interviewee asks DVA what time it is	2
Asking for weather forecasts	Interviewee requests weather forecast from DVA	16
Information seeking	Interviewee requests various kinds of information (e.g. facts, numbers, meaning of words)	34

(IV) Children's probing schemes when exploring DVAs' intelligence and identity

Code name	Code description & Examples	References
Probing DVAs' knowledge		
Interpersonal knowledge	Interviewee asks questions related to knowledge of interpersonal nature (e.g. "Alexa, what's my father's name?")	
Nominal facts	Interviewee asks questions that can be answered with one word or sentence (e.g. "Alexa, do polar bears live at the South Pole?")	21
Open information	Interviewee asks open questions that require more than one word or sentence (e.g. "Alexa, tell me something about America!")	5
Meaning of words	Interviewee asks for the meaning of certain terms of names (e.g. Jordan: "Alexa, what does Jordan mean?")	6
Probing skills		
Foreign language skills	Interviewee probes DVAs' foreign language skills (e.g. "Hey Google, can you say a word in Spanish?")	6
Mathematical skills	Interviewee probes DVAs' mathematical skills (e.g. "Hey Google, what is 2000 + 2000?")	13
Generic skills	Interviewee probes DVAs' generic skills (e.g. "Hey Google, say a word!")	5

Probing DVAs perceived identity	Interviewee asks questions related to DVAs' 'personality' traits or 'personal' background information (e.g. name, age, relationship status)	15
Probing DVAs reflexivity	Interviewee asks questions that would require introspection when answered by a human (e.g. "Hey Google, how are you doing?" or "Alexa, where are we right now?")	22

(V) DVAs' response behaviours perceived as human behaviour (focus group session 2)

Code name	Code description & Examples	References
Dynamics of DVAs' response behaviour		
Delayed responses	Interviewee perceives DVA's slow response behaviour (e.g. long gaps between voice command and response, or several voice commands necessary to obtain desired response) as an indicator for the presence of human control over the device	9
Permanent responsiveness	Interviewee perceives observation that DVA always reacts to voice commands with verbal responses as an indicator for the presence of human control over the device	2
Non-responsiveness	Interviewee perceives observation that DVA does not always react to voice commands with verbal responses (e.g. although lights go on after wake word, no verbal response is provided) as an indicator for the presence of human control over the device	5
Nature of responses		
Creative responses	Interviewee perceives rather original and imaginative DVA-responses (e.g. puns or very playfully formulated answers) as an indicator for the presence of human control over the device	8
Response mistakes	Interviewee perceives objective mistakes or inconsistencies made by the DVA (e.g. Google Assistant reacts to the wake word 'Alexa') as an indicator for the presence of human control over device	2

(VI) DVAs' response behaviours interpreted as machine behaviour (focus group session 2)

Code name	Code description & Examples	References
DVAs' conversational qualities		
Lack of common sense	Interviewee perceives DVA's struggle to respond to questions that would normally require common sense if answered by a human (e.g. describing how a human being looks like) as an indicator for actual machine behaviour (i.e. absence of active human control)	5
Lack of understanding	Interviewee perceives DVA's struggle to understand questions in the intended way as an indicator for actual machine behaviour (i.e. absence of active human control)	6
Dynamics of DVAs' response behaviour		
Immediate provision of nominal facts or quantities	Interviewee perceives DVA's capacity to provide nominal facts or quantities instantly after a question was asked as an indicator for actual machine behaviour (i.e. absence of active human control)	7
Permanent responsiveness	Interviewee perceives that fact that DVA always reacts to voice commands with verbal responses as an indicator for actual machine behaviour (i.e. absence of active human control)	2
Promptness	Interviewee perceives DVA's instantaneous response behaviour (e.g. very short gap between voice command and response) as an indicator for actual machine behaviour (i.e. absence of active human control)	6
Non-responsiveness	Interviewee perceives observation that DVA does not always react to voice commands with verbal responses (e.g. although lights go on after wake	2

	word, no verbal response is provided) as an indicator for actual machine behaviour (i.e. absence of active human control)	
Nature of responses		
Accuracy of responses	Interviewee perceives accuracy of DVA-responses (e.g. providing exact quantities in response to a question) as an indicator for actual machine behaviour (i.e. absence of active human control)	4
Familiarity with standardised responses	Interviewee expresses familiarity with the exact wordings of standardised DVA responses, especially standardised failure response (e.g. "I do not have an answer, unfortunately!") and perceives this as an indicator for actual machine behaviour (i.e. absence of active human control)	5

Notes. Tables shows detailed code book of the inductive thematic analysis that was applied to the focus group data (i.e. transcripts as well as audio-visual recordings). Each row represents a node within a theme, including the node's name, its description, examples as well as its frequency in the data. The analysis was conducted in NVivo 12.

Statement in Lieu of an Oath

I hereby confirm that I have written this thesis on my own; I have not used any other media or materials than the ones referred to in this thesis. s

This thesis has not been submitted to any other examination office and it has not been published yet.

Oxford, 8th August 2019

