

Journal of • Virtual Worlds Research

jvwresearch.org ISSN: 1941-8477

Volume 3, Issue 2
December 2010
Virtual Worlds for Kids



Volume 3, Number 2

Virtual Worlds for Kids

December 2010

Editor-in-Chief

Jeremiah Spence

Guest Editors

Sun Sun Lim
National University of Singapore

Lynn Schofield Clark
University of Denver, USA

Cover illustrator

Lim Su Pei
National University of Singapore

Editorial Assistants

Cao Yuanying
National University of Singapore

Shobha Vadrevu
National University of Singapore

Reviewers: Denise Anthony, Iccha Basnyat, Anne M. Burke, Leanne Chang, Pin Sym Foong, Sara Grimes, Elizabeth Hayes, Jackie Marsh, Alex Mitchell, Elmie Nekmat, Alvin Saw Teong Chin, Becky Herr Stephenson, Shobha Vadrevu, Karen Wohlwend, Joshua Wong Wei-Ern



The Journal of Virtual Worlds Research is owned and published by the Virtual Worlds Institute, Inc. – Austin, Texas, USA. The JVWR is an academic journal. As such, it is dedicated to the open exchange of information. For this reason, JVWR is freely available to individuals and institutions. Copies of this journal or articles in this journal may be distributed for research or educational purposes only free of charge and without permission. However, the JVWR does not grant permission for use of any content in advertisements or advertising supplements or in any manner that would imply an endorsement of any product or service. All uses beyond research or educational purposes require the written permission of the JVWR. Authors who publish in the Journal of Virtual Worlds Research will release their articles under the Creative Commons Attribution No Derivative Works 3.0 United States (cc-by-nd) license. The Journal of Virtual Worlds Research is funded by its sponsors and contributions from readers. If this material is useful.

**Volume 3, Number 2
Virtual Worlds for Kids
December 2010**

Who's Watching Your Kids?

Safety and Surveillance in Virtual Worlds for Children

By Eric M. Meyers

School of Library, Archival and Information Studies, University of British Columbia

Lisa P. Nathan

School of Library, Archival and Information Studies, University of British Columbia

Kristene Unsworth

The Information School, Drexel University

Abstract

Virtual playgrounds designed for children 4-12 years-old are among the fastest growing segments of the Internet. These spaces offer the promise of new social, educational, and creative opportunities for young people. These opportunities, however, are associated with new risks due to the nature of online interactions. This article presents a socio-technical investigation of two virtual worlds for children. Informed by Value Sensitive Design, the work highlights the tensions between supporting age appropriate, developmental play and constraining inappropriate behavior to ensure "safety." This article concludes with implications for the design of virtual spaces for young people. We note in particular that the features of children's worlds are affected by their framing, specifically by the language and practices that are encouraged by the interaction design. We challenge designers to consider how children may interpret technical feature in practice, and the long-term implications of features intended to keep children safe.

Keywords: children's virtual worlds, values, value sensitive design, socio-technical

Who's Watching Your Kids?

Safety and Surveillance in Virtual Worlds for Children

“The avatar seemed for awhile the very apotheosis of the American dream, and the antithesis of the panopticon: it embodied the fantasy that we can refashion our selves in any manner we wish. We can't be under surveillance if no one knows who we are: what we look like, what color we are, what religion and ethnicity, what sex.”
(Winokur, 2003)

Cyber-stalking. Cyber-bullying. Cyber-tattling. These anxiety-inducing terms have become part of the conversation around children's online interactions. As children increasingly move toward immersive spaces online, the promise of the avatar as emancipating technology has not been fulfilled. In contrast to this utopian vision of perfect privacy and refashionable identity, youth-focused virtual worlds have been received with a mixture of optimism and moral panic. Scholars have identified the potential of these spaces for literacy (Merchant, 2009), learning (Beals & Bers, 2009) social interaction and creativity (Thomas & Brown, 2009); others have guardedly observed that these spaces carry opportunities as well as risks (Harwood & Asal, 2007; Livingstone, 2009; Livingstone & Brake, 2010; Palfrey & Gasser, 2008). Articles in the popular press, particularly when these spaces first emerged, tended to focus on risks: challenges of child safety and privacy (e.g. Flanagan, 2007), pro-social development and consumerism (e.g. Slatalla, 2007), cheating and deviant behavior (e.g. Benderoff, 2007).

And yet, despite a mixed reception among parents, educators, and service providers; virtual worlds for children continue to grow in number and popularity. While adult virtual world use has stagnated, there are dozens of virtual worlds for kids ages 5-18, and as many more under development (KZero, 2009). It is estimated that 10 million children use virtual worlds on a regular basis, and it is estimated that 54% of children online in the U.S. will visit a virtual world every month by 2013 (Virtual World News, 2009). One factor that may hinder the development of youth-focused virtual worlds is the challenge and expense of monitoring them, both to safeguard their users and to enforce behavioral norms (Economist, 2009).

To date, little research has investigated how virtual worlds for children manage the inherent tension between supporting age-appropriate play and constraining inappropriate behavior to keep children “safe”. Nascent empirical work in this area (Marsh, 2010; Meyers, 2009)

suggests that sites like Club Penguin, Woogi World, BarbieGirls and Pixie Hollow are designed to be safe places for children to play, but each site in this genre appears to have different interpretations of what is necessary to keep a child safe online. These interpretations are instantiated in specific design features, such as monitored chat systems, typically targeted to the concerns of parents rather than the intended users of the technology, the children. In some cases, children are recruited to monitor each other to enforce the site's terms of service. Such design decisions have both social and moral development implications.

This paper presents a values-oriented, socio-technical analysis of these hugely popular virtual spaces and how they support and constrain the values of different stakeholders. Specifically, we investigate two virtual worlds (Woogi World and Club Penguin) with a focus on the sites' technical features and behavioral cues designed to support value concerns (e.g. privacy, safety, security) and critically examine the potential influence of these features from different stakeholder perspectives. The worlds discussed above were selected as they represent two different approaches to managing the safety of children, as will be explained later in the paper.

We address the following questions in our work:

- How do virtual worlds for children (explicitly and implicitly) define safety?
- How is this definition of safety instantiated in technical features and behavioral cues that scaffold children's socio-technical practice?
- What is the role of surveillance in the design of safe virtual environments for children?

The work reported here draws on the tripartite methodology of Value Sensitive Design (VSD) (Friedman, Kahn, & Borning, 2006) to conduct investigations of two children's virtual worlds: Club Penguin and Woogi World. Our process involves two integral investigations: 1) a conceptual investigation, where we analyze different stakeholders such as children, adults, and designers, their roles in using the virtual world, and tensions that might emerge in use; and 2) a technical investigation, in which we focus on specific technical features these worlds employ, such as moderated chat systems, and the practices associated with those features. Taken together, the analysis from the conceptual and technical investigations informs the development of brief case studies that synthesize the findings from the VSD-informed investigations. The following discussion teases apart how subtle differences in design result in a different ethos and experience for the site user. The article

concludes by outlining an empirical investigation informed by the findings and a discussion of potential implications for the designers of children’s virtual worlds.

The project follows in the tradition of earlier conceptual explorations of values in teen game design (Flanagan, Howe, & Nissenbaum, 2005; Flanagan, Belman, Nissenbaum, & Diamond, 2007). Although the analysis is influenced by contemporary critiques of privacy and surveillance (Andrejevic, 2007; Lyon, 2001), its primary aim is to scaffold future empirical research into the long-term, systemic influence of virtual world engagement on children’s development (Nathan, Friedman, Klasnja, Kane, & Miller, 2008). The values orientation of the work distinguishes this project from others rooted in literacy and communication theory.

Motivation

While virtual worlds for young people have been available since the 1990s, their popularity and use in the home environment has exploded only in recent years. To date, there are few research programs examining virtual worlds for children less than twelve years of age. The existing published research has focused on developing frameworks to understand children’s experiences in these spaces from the standpoint of play (Marsh, 2010), learning and cognitive development (Beals & Bers, 2009), and information seeking and use (Meyers, 2009). Marsh’s interviews with ten young people examine the relationship between online and offline play in virtual worlds, identifying that sites like Club Penguin and Barbie Girls constitute “friendship-driven” rather than “interest driven” spaces (Ito, Baumer, Bittanti, boyd, Cody, Herr et al., 2008). Beals and Bers describe how developmental theory might inform the design of children’s worlds, but offer only high-level guidance for the creation of these spaces, noting for example that children’s spaces need rules, creative outlets, and facilities that permit communication. Meyers identifies the “information work” associated with virtual worlds, and suggests that the most creative practices may be found outside these worlds in affinity spaces. These framings are valuable contributions to the conversation about children’s immersive technologies, and all conclude that these spaces represent an emerging research area with potential benefits for young people. Thus far the research literature has focused more on the use and mediation of these sites than on the design and evaluation of technical features, or how these sites instantiate stakeholder values.

Virtual worlds offer the promise of new social, educational, and creative opportunities

for young people. These opportunities, however, are associated with new risks due to the nature of online interactions (Livingstone & Brake, 2010). The Federal Trade Commission (2009) recently released its report to the United States Congress, documenting that a number of youth-oriented virtual worlds contained violent or sexually explicit materials, predominantly in the form of text chat. The report suggests that a combination of technical features, parent oversight, and user education is necessary to manage the risks of virtual worlds for youth. The challenge in our current childhood landscape is that any risk is perceived as unacceptable. We appear to live in a risk-averse culture that works not just to reduce unnecessary risk, but also to eliminate the slightest possibility of harm to children (Byron, 2010). The ability to identify and manage harmful or threatening situations, however, is a key developmental capacity, and allows youth to manage uncertainty as they mature (Christensen & Mikkelsen, 2008). Efforts to eliminate exposure to risk may limit opportunities for young people to develop essential cognitive and metacognitive strategies (Livingstone, 2009). Donovan and Katz (2009) argue that children are informed social actors and that attempts by adults to monitor their online behavior and filter their websites has a chilling effect on children’s and youth’s privacy and free speech.

Speaking more broadly, there are important societal motivations for better understanding the competing values on digital playgrounds. First, as these sites grow in popularity and number, designers, policy makers, and parents need guidance in creating, regulating, and facilitating children’s online environments. The adults mediating these worlds did not grow up in immersive spaces; children are venturing into largely uncharted territory. Second, these worlds provide us an opportunity to think critically about societal expectations of online environments. Is Club Penguin (owned by Disney Interactive) the equivalent of a sandbox at a public park, where parents and other caregivers are expected to mediate children’s interactions to teach socially appropriate and “safe” behaviors? In this framing, some level of risk is assumed by those who take their children to the sandbox. Alternatively are these sites like a ride at Disneyland, where movements and interactions are tightly constrained and monitored. In this framing, the parents, who watch as their child is buckled into the ride, assume almost no risk and the Disney Corporation is expected to prevent any harm from happening to the children. These issues deserve explicit attention and deliberation. The answers that emerge will affect the future of digital play.

Value Sensitive Design

Developed by Batya Friedman’s research team at the University of Washington, Value Sensitive Design (VSD) provides a principled approach to the analysis of interactive system design (Friedman, 2004; Friedman & Kahn, 2003; Friedman et al., 2006). Value sensitive design (VSD) represents a methodological practice based upon an interactional theory which together provide a framework for scaffolding the critical examination of human values during the design, use and evaluation of such systems. The VSD approach has been successfully used in diverse projects including computer simulations in urban planning (Borning, Friedman, Davis, & Lin, 2005), the design of code sharing software (Miller, Friedman, Jancke, & Gill, 2007), the creation of end-user license agreements (Friedman, Smith, Kahn, Consalvo, & Selawski, 2006), and the analysis of technology practices in intentional communities (Nathan, 2009).

Friedman and her colleagues contend that through interactions with information systems, values can be reinforced and/or destabilized. The values that VSD research focuses upon are those that are asserted to be of moral weight. However throughout the broad range of VSD investigations, other values such as democratic fairness and usability have been investigated (Friedman, 2004). Explications of VSD state that they define a value as “what a person or group of people consider important in life” (Friedman et al., 2006, p. 2). Examples of human values that may be integral to understanding the relationship between youth and interactive systems like a virtual world include self-actualization, privacy, autonomy, community, safety, and anonymity.

Values are uncovered and conceptualized through three types of rigorous, iterative and integrative investigations that are labeled conceptual, empirical, and technical. There is no strict ordering of these investigations and they can be undergone simultaneously. In the ideal manifestation of a VSD project, the investigations happen in such a way that they are able to inform each other, making each investigation more robust. However, it can be unmanageable, especially within the time and resource constraints of most research projects, to conduct all three investigations concurrently. A brief description of each type of investigation is provided below.

- Conceptual investigations typically begin by identifying the stakeholders who will be affected by the design under study. This would include both the direct stakeholders, those who will actually use the device and indirect stakeholders, those who may not physically touch the technology, but whose lives will be influenced through a direct

stakeholder’s use of the technology. For example, a conceptual investigation of a satellite monitoring system designed to assist law enforcement personnel of a small community would identify the community’s police officers as direct stakeholders and individuals living in the community, but without direct access to the system, as indirect stakeholders. Identifying and defining the values implicated by a technology is the second key component of a conceptual investigation. The example of the satellite monitoring system may implicate values such as privacy and personal safety. In a conceptual investigation, the work of philosophers and other scholars is drawn upon to develop robust understandings of the values under scrutiny.

- Technical investigations are primarily concerned with the features of the technology under investigation and may include actually designing a technology or investigating how specific features implicate certain values. A technical investigation focuses on the technology itself. However, we recognize that what a technical feature “does” is socially constructed and negotiated between site users and site developers. Thus, the technical investigation reported on here takes an explicit look at the practices that develop to help direct stakeholders understand and manage technical features.
- Empirical investigations serve to expand the conceptual investigations by using both language and observable actions (through surveys, questionnaires, interviews, participant observations, experiments) to deepen the investigators’ understanding of both the context and the values implicated through the interactions with a design. An empirical investigation focuses on the actions of human actors.

Methodology and Findings

To deepen our understanding of children’s virtual worlds we conducted VSD-informed conceptual and technical investigations of two such sites: Club Penguin (<http://www.clubpenguin.com>) and Woogi World (<http://www.woogiworld.com>). Both spaces are designed expressly for youth, but each approaches the values of safety and surveillance differently. These two sites were selected from a range of virtual worlds examined (others include EcoBuddies, BarbieGirls, Webkinz, and HandiPoints, along with several worlds aimed at older children, such as Gaia, Habbo, Millberry), and identified by their characteristics as contrasting cases. Material for both investigations was drawn from technical and documentary features of the sites, as well as investigators’ interactions with the

sites as users (2007-2010). Investigators also identified and interacted with supplementary online materials related to these virtual worlds including user-generated content (e.g. blogs and websites, YouTube channels), review sites (e.g. Commonsensemedia.org), and website analytics. To be clear, the analyses presented below are illustrative not exhaustive. There are myriad technical features and practices that could be presented in depth, but this analysis emphasizes those features that support or constrain stakeholders’ perceptions of safety and surveillance.

The findings within this section are organized as follows: first, we provide highlights related to stakeholders, values and interaction that arose from the conceptual investigations of both sites. We follow-up the highlights from the conceptual work with findings of note from the technical investigation. An organizational table helps to draw apart technical features from the behavioral cues that the sites put in place to frame how users interpret and interact with the technical features. To assist in synthesizing the findings from the two investigations, the section concludes with two case studies that demonstrate how surprisingly similar technical features are interpreted through the different behavioral expectations of the sites to create two disparate interpretations of a “safe” virtual playground.

Conceptual Investigation

Due to space constraints we focus on specific insights afforded by four elements of our conceptual investigation into children’s virtual worlds, namely: 1) Multiple stakeholders; 2) implicated values; 3) value tensions; and 4) systemic interactions.

Multiple Stakeholder Roles

While virtual worlds such as Club Penguin and Woogi World are designed expressly for younger users, there are often multiple stakeholder roles, direct and indirect, which are affected by use of these technologies. In the VSD tradition, direct stakeholders are those who interact with the technology, such as the children who login to the world of Club Penguin. The parents of these children fall into the role of indirect stakeholders if they never login to the system themselves. Indirect stakeholders do not interact with a technology, but they are still influenced by others’ use of the system. Even though a parent does not login to Club Penguin, he may be affected by the time his daughter spends on the computer caring for her

penguin and socializing with friends. The VSD methodology also recognizes that stakeholders may take on more than one role with regard to a given technology. Parents may begin as direct stakeholders, initially logging in to a site to “check it out”. However, after that initial foray, parents may passively slide into the role of indirect stakeholder if they never log in to the site again. VSD work also reminds us to be aware of nefarious roles that users may take on (Nathan et al., 2008). Kids may be users and beneficiaries of a virtual world; at the same time acting maliciously or subversively to undermine the experience for other users through hacks or cheats. The boundaries between these roles (user vs. hacker, player vs. cheat) may be fluid and dependent on the norms of behavior in a technological space, or the expectations of other direct and indirect stakeholders.

Implicated Values: Safety, Surveillance and Privacy

Safety, according to the Oxford English Dictionary (OED), is a term that defines “the condition of being safe,” and safe, “protected from or not exposed to danger or risk” (OED, pp. 1264, 1265). Safety awareness initiatives and security policies that aim to provide some type of monitoring for children’s online behavior echo concerns regarding risk in childhood in general (Staksrud & Livingstone, 2009). Monitoring children with technology throughout their day has become commonplace in the U.S. and other parts of the developed world. Children are watched in the crib via “baby monitors” and teddy bear cameras, in the hallways of schools and shopping malls through closed-circuit television systems, and online with the assistance of software applications such as Net Nanny and CyberPatrol. Such monitoring is rooted in adult concerns for the safety of children, whose sense of danger and appropriate behavior is still developing.

Beyond managing the risks of growing up, the everyday behaviors of adults and children are coming under increased surveillance under the auspices of personal safety, corporate risk management, and national security. Such moves increasingly equate “safety” with reduced personal privacy and increased surveillance, either by corporations, the state, or organizations acting on the state’s behalf. We can define surveillance as the “collection and processing of personal data, whether identifiable or not, for the purpose of influencing or managing those whose data have been garnered” (Lyon, 2001, p. 2). In the United States, the surveillance of activity through data collection comes with the limited right of the user or citizen to know what data is gathered, and how it is used to safeguard the individual

(Andrejevic, 2007). In virtual worlds and other social software, end-user license agreements (EULAs) or Terms of Service (ToS) are written broadly to permit the hosting company to gather and save chat transcripts and other user information indefinitely to improve “the Service” (e.g. Linden Research, 2010), ostensibly for the benefit of the user.

Lyon (2001) writes that there are two faces to surveillance: care and control. When we use technology to watch over children, we may have the primary instinct to reduce their privacy in order to keep them from physical or psychological harm. We may also wish to observe children’s activities to ensure that they do not get into mischief. This latter purpose, though not always in conflict with the former, applies a different set of criteria to the situation. A different set of values is at play when we prioritize behavioral outcomes through surveillance. The developers of virtual worlds for children appear to have several different goals in mind when designing features to keep children “safe”: providing fun and entertainment for children, pleasing adult gatekeepers, avoiding litigious situations that may emerge through use, and gaining (and maintaining) profitability. While these goals are not necessarily at odds with each other, an emphasis on one may constrain the ability of the virtual world to fulfill another. This begs the question of whether monitoring and surveillance activities in virtual worlds for children privilege control over care.

Value Tensions

Previous VSD investigations call attention to the design challenges that emerge when values are in tension with each other (e.g., Miller et al., 2007). Sensitivity to value tensions helps us identify when different stakeholder roles bring different views and values to their interactions with a given system. A system may actively support a value of a key indirect stakeholder (e.g., parents and safety) while constraining a value of a direct stakeholder (children and self expression). We can see how this plays out in the different interpretations children and their parents bring to virtual worlds. The features of children’s virtual worlds, like chat systems with limited dictionaries or active monitoring, can impact different stakeholders in different ways. Limited chat supports the values of one key stakeholder (e.g. parents who value safe, appropriate conversations among youth) for the purpose of constraining the values of another stakeholder (e.g. pedophiles who seek freedom of expression to groom young children) while limiting the activities of another stakeholder (children who wish for freedom of expression to try on different identities). These tensions

often represent design trade-offs, where supporting a range of values may be contradictory or impossible. By uncovering these tensions, designers can envision the long-term effect of design decisions, and recognize when supporting one set of values over another (e.g. supporting parents’ concern for safety and appropriate communication over children’s communicative rights) may result in unintended use scenarios, user work-arounds and hacks, or undesirable activity in other aspects of the system (Nathan et al., 2008).

Systemic Interactions

Recent VSD work highlights the systemic and emergent effects of interactions with interactive information systems (Friedman & Nathan, 2010). Similar to offline geographically placed environmental ecologies, online virtual world ecologies are made up of countless features that interact and influence each other over time. A feature of a virtual space does not exist in isolation; it influences other system features, users, and activities outside the system. Participation in virtual spaces is consequential: the user’s presence is part of the experience, and consequential to the way other users experience the space (Pearce, 2009). As a result, virtual worlds for children, through wide-scale participatory structures, have systemic effects and emergent qualities that support or contradict the intentions of the system designers. The communicative limits of children’s virtual worlds, for example, have effects within the virtual world and outside of it: children incorporate alternate chat vocabularies to circumvent word lists, construct blogs and videos to convey hacks and cheats, and actively disseminate information about the world that affects the way other users experience and perceive it.

Technical Investigation

The conceptual investigation highlighted the issues that emerge in children’s virtual worlds surrounding safety and surveillance. The technical investigation illustrates these tensions and effects by examining two example virtual worlds in-depth. These two spaces were chosen to provide two contrasting cases of safety and surveillance. The two worlds have instantiated different choices in terms of technical features and behavioral cues that in turn shape and illuminate the ethos of the sites themselves. Table 1 highlights just a few differences and similarities between these two virtual worlds in terms of these choices. For each world we identify the premise that guides user interactions, how the site defines safety

in its information materials, technical and supporting features.

	CLUB PENGUIN	WOOGI WORLD
Slogan	“Waddle around and meet new friends”	“Teaching kids Internet safety, life values, and fun!”
Premise	Anthropomorphic penguins populate an animated snow-covered community of shops, games, and activity spaces. Children register with the site to engage in activities using the penguins as avatars. Children are encouraged to “waddle around” and play independently or with others.	The Woogis are an alien race corrupted by excessive online time. The children of Earth help the Woogis by adopting them and teaching them appropriate activities and limits. Children control a Woogi character, perform tasks, play games, and earn virtual currency.
Definition of Safety	Club Penguin explicitly and prominently addresses safety concerns with a description on the parent page of chat limitations. <i>“Ultimate Safe Chat limits what users can say to a predefined menu of greetings, questions and statements, as well as emotes, actions and greeting cards.”</i>	Woogi World includes a definition of safety in its parents’ section: <i>“A combination of appropriate mindset and behavior validated by a community of like-minded people.”</i>
Technical Feature: Account creation	Email address required to activate account; No personal information required to create user ID aside from email address; Accounts require regular use to remain active.	Email address required to activate account; No personal information required to create user ID aside from email address.
Technical Feature: Chat	Multi-level chat including “ultimate safe chat” option; can chat with or “friend” anyone in the “room”; Chat does not support distribution of files or hyperlinks; Sharing personal information (age, phone number, address) is not permitted.	Users must complete a training program on appropriate and inappropriate chat behavior; Honors users have special chat privileges; Multi-level chat including “safe” chat; can chat with or “friend” anyone in the “room”; Chat does not support distribution of files or hyperlinks; Sharing personal information (age, phone number, address) is not

		permitted .
Technical Feature: User Profile	User profiles contain no personal information aside from login name.	
Technical Feature: User-generated content	All content is generated in-world or by site designers.	
Technical Feature: Activity Monitoring	Automated monitoring of chat ¹ ; Active adult & peer monitoring of behavior; Ability to report other users for inappropriate behavior.	Automated monitoring of chat; Active adult & peer monitoring of behavior; Ability to report other users for inappropriate behavior.
Other Supporting Features	Parent Section; Parent reporting options (for time on site, but not activities); Special socialization features limited to paying members	Parent section; Big Wig blog; Premise supports limited online time; Special socialization features limited to paying members

Table 1. Site comparison.

Through our technical analysis, we identified a series of site features that potentially affect the safety and privacy of users, either explicitly or implicitly. Explicit technical features include the account creation and user profile, chat features, user content generation, and activity monitoring. Implicit features include the site slogan and premise, as well as how the site defines safety, explicitly or implicitly. The slogan and premise of each world, although not obvious technical features of the site, set the tone of the site, and perhaps more importantly, frame the technical features for child and adult stakeholders. This provides an ethos that shapes stakeholders’ emergent practice. For example, the Woogi World slogan, “Teaching kids life values, Internet safety, and fun,” makes explicit that educating users in safe Internet practices is a goal of system. It arguably addresses parents or adult guardians rather than youth in that it speaks of “kids” in the third person.

Many of the site features are held in common, suggesting that different sites use a common template of technical features for user safety and privacy (“safety by design”). Both sites analyzed here require users to provide a valid email address, but do not validate the age of the user. Instead, nearly all personal information, including age and location, are kept from other users. In the absence of being able to confirm that all users are children, these

¹ certain words are automatically removed from chat strings or replaced with symbols/ other characters

sites keep all users at the same high level of anonymity.

Differences largely reside in the ethos of the virtual world, the textual and pictorial framing the sites use to balance the technical features with certain social interactions and practice. Both regard inappropriate text chat as a key danger to users; this is consistent with other virtual worlds and social networks designed for younger users. Most of the textual information about the spaces is directed at adults in the roles of parents, teachers, or caregivers as key stakeholders in the success of these playgrounds. Club Penguin relies heavily on its technical affordances to police chat, relying less on human monitoring, either through adults or peers, to eliminate inappropriate sharing of personal information (names, addresses, phone numbers) or sexually explicit and crude/mean/harassing language. This may be due to the enormous traffic through this space, which would be difficult to monitor effectively through human moderation. Woogi World, on the other hand, employs significant user education and social pressure to encourage users to conform to conduct expectations, in addition to many of the same technical features, as a first line of defense against inappropriate content. Other virtual worlds employ different blends of these two mechanisms to achieve a “safe” environment. All the virtual worlds these authors are aware of in this genre include some kind of surveillance mechanism that is beyond the oversight children would experience on a physical playground, and their privacy in these spaces is significantly curtailed at the expense of regulating speech and behavior.

“Safety” in children’s virtual worlds never means freedom from commercial inducements and hard-sell advertising, however. Subscription levels – paying members vs. registered users – strongly affect how users interact in these spaces. The degree of freedom permitted to paying members of these and other virtual playgrounds, including chat privileges, social organizations, and displays of status and belonging, is at a significantly higher level. Membership has its privileges, and virtual worlds are designed to showcase this disparity: the haves and have-nots are identifiable at a glance by the appearance of their avatars. In some ways, this makes the non-paying members not only inferior but also suspect, and adds to the ethos of lateral (peer-to-peer) surveillance that is layered on top of the constant adult monitoring.

Emergent Interactions: Case Example of Children’s Virtual Worlds

We present the following case examples to illustrate how these two virtual worlds

frame parent and child expectations of safe behavior in online virtual environments. The two cases explore the interactions that emerge as users interact with the technical features and the narratives that surround these features.

Case 1: Club Penguin

Club Penguin is a virtual playground employing the metaphor of the Antarctic where children are encouraged to “waddle around and make new friends.” Users register with the site to engage in activities as anthropomorphic penguins in an animated snow-covered community of shops, games, and activity spaces. Users can perform tasks in the community (making pizzas at the pizzeria, unloading coffee beans at the café, and of course, rounding up puffles) in order to earn virtual currency (“coins”). They use this currency to purchase clothing, pets, and accessories for their penguin avatars and to decorate their “igloo” or private online space. A large number of features, including the ability to buy most clothing items, are reserved for subscription users who pay \$5.95 per month (as of August 30, 2010). Users can create friend lists and chat with other penguins via a simple dashboard interface. Adult moderators monitor chat, and can penalize through temporary or permanent banishment users who violate the chat guidelines, namely “mean” or age-inappropriate talk, such as sex or drug references.

Although an explicit definition of safety is not provided, Club Penguin prominently addresses safety concerns with a description on the parent page of chat limitations. “Ultimate Safe Chat limits what users can say to a predefined menu of greetings, questions and statements, as well as emotes, actions and greeting cards. When it comes to chatting, these users can only see other Ultimate Safe Chat messages” (Club Penguin, 2009). Thus, within Club Penguin’s Ultimate Safe Chat, the value of safety is threatened by conversations in which either children or adults reveal inappropriate information. Club Penguin is an all-ages site, even if it is designed for children. There is no age verification, and parents are encouraged to interact with their children on the site. This suggests the space is open to the possibility of contact between and among users of any age, albeit anonymously and in the form of a penguin avatar. Unable to truly control who is in the space, safety is “realized” through technical features that restrict conversation among users.

This is not to suggest that inappropriate behavior cannot happen in these spaces. Users are constantly testing the bounds of what can and cannot be said and done, despite (or

perhaps because of) the technical features. The restriction on sharing phone numbers and addresses is easily circumvented with alternate words considered “safe” by the chat filter (e.g. won too tree for fakes tweet = 1234 Fake Street). Witness the plethora of YouTube videos documenting cussing hacks (e.g. “fuckity” <http://www.youtube.com/watch?v=mHO6BTUbmYw>). Users also engage in explicit pseudo-sexual play (“pseudo” in that there is only so much one can do with an asexual cartoon penguin) using code phrases developed outside the restrictions of the space, such as the proposition “wanna”. Contributors on CommonSenseMedia.org, a site that provides user-generated reviews of popular children’s movies, games, and television shows, describe the various work-arounds they have used or witnessed, despite generally positive ratings.

To buttress the technical limitations of the safety features, users are recruited to aid in making Club Penguin safe through active peer monitoring. Young children often engage in gender policing (“boys don’t wear pink”), and this tendency is leveraged to police conversation in virtual spaces as well. After 30 days in Club Penguin, users may apply for Secret Agent status, which permits a user to tattle on offensive peers. Innocuous schoolyard taunts (“Boys rule, girls drool!”) may be treated with a barrage of text rebukes and entreaties to the monitor to censure the offending penguin. By assigning the coveted title of Secret Agent to tattlers, paired with the anonymous nature of monitoring and peer surveillance in this space, Club Penguin has established a system through which any conversation that offends or is interpreted as “mean” by another user may be reported and penalized.

Case 2: Woogi World

According to the storyline behind Woogi World, the Woogis are an advanced alien race corrupted by excessive online time without careful monitoring. The children of Earth can help the Woogis by adopting them, and teaching them appropriate online activities and behavioral limits. In Woogi World, children adopt a Woogi character (control its movements), perform tasks, play games, and earn virtual currency (“Woogi Watts”), which they can use to decorate their “wigwams” or accessorize their adopted Woogi avatars. Virtual activity spaces are learning oriented, and include a science lab, savings bank, and a library. Children are also provided incentives for off-line activities (e.g., helping with chores, playing board games with family members). A simple, adult monitored chat interface is provided, and several games are designed as social, multiplayer experiences. Woogi World describes its goals on the site’s front page: “teaching kids Internet safety, life values, and fun.”

Woogi World includes a definition of safety in its parents section: “A combination of appropriate mindset and behavior validated by a community of like-minded people” (Clare, 2008). Woogi World, however, does not describe specific technical features that promote safety. Instead it describes to parents how the site’s rules, standards, and regulations are embodied by the children who are now serving as their own moderators, in addition to the site’s adult moderators. In order to enable chat features, Woogis (users) have to go through a short, intense training program that asks them to correctly distinguish between appropriate and inappropriate chat phrases. In Woogi World, as in Club Penguin, the value of safety is threatened by conversations in which inappropriate information is shared. Woogi World is also an open, all-ages space. However, Woogi World promotes habits of mind in games and activities. For example, in the “Service Station”, which looks like a gasoline pump, users document the good deeds they have done for others in order to advance their status in the system. Thus, in addition to technical features, Woogi World addresses parents’ concerns for safety through developing a strong ethos of “appropriate” interactions.



Figure 1. The Big Wig.

While Club Penguin includes several non-player characters that appear from time to time for the purpose of introducing new activities (Rockhopper the Pirate, Aunt Arctic, the dojo Sensei), the key non-player in Woogi World, the Big Wig, acts as a boss character or

leader, actively promoting the site’s value system, cautioning users, and providing information that helps mediate play. The Big Wig (see Figure 1) looks remarkably like a cartoonish Big Brother from George Orwell’s 1984, appearing as a giant head with a prominent mustache. He extorts users to sign-up for Woogi Honors, the fee-based aspect of the virtual world that bestows users with additional rights and privileges – “get COOL things that regular Woogis don’t” – such as more advanced vocabulary in chat, pet ownership, clothing options, and the ability to create and join social groups. In addition, by joining Woogi Honors, users may join clubs and groups that provide another level of socialization within the space. One such club, the WW Police, has as its motto: “we are going to find all bad woogis in woogi world [sic]”. The Big Wig regularly reminds users that their subscriptions will make Woogi World even better.

Discussion: Implications for Design

How do we keep young Internet users safe? Whose responsibility is it to keep these children safe? As many scholars and policy makers suggest, it requires a mix of user education, improved technology, and parental oversight. One of the challenges is finding the right mix of these elements, and how to balance social and technical features while envisioning their systemic effects. Rather than responding to moral panics or isolated breakdowns, designers and policy makers need to think holistically about how virtual worlds – virtual communities of impressionable users in this case – and design decisions influence both specific behaviors as well as perceptions of how online interactions work. If a virtual world like Club Penguin is a child’s first foray into online social networking and communication, these perceptions may have lasting impact.

A key issue identified through this analysis exists between children as direct stakeholders, who value fun, sociality and communication, and the adult stakeholders who provide access to these spaces, who value safety. The designers of Woogi World, recognizing parents as a key gatekeeper, have emphasized the parents’ values over that of the children, and instantiated this value in features and social practices that emphasize limited communication and a culture of tattling. Club Penguin has similar technical features, but is geared more for children, but the social pressures focus on children as consumers.

Although they take a different orientation to these two stakeholder groups, both worlds use surveillance as both a protective as well as a disciplinary tool. Teaching children

moral and ethical behavior is important whether they are citizens of virtual or physical worlds. Virtual worlds for children provide diverse means of modeling appropriate behavior, but some of the systemic effects of an emphasis on safety can be troubling. Encouraging an anonymous, politically correct tattle-tale culture and permitting hyper-vigilance to evolve into vigilantism to enforce social norms sends an unfortunate message to young users. Furthermore, wrapping model behavior in an aggressive commercialism suggests to users that purchasing citizenship is an effective and rewarding form of participation.

From this analysis, we propose three implications for the design of virtual worlds for children. First, we suggest that designers need to be cognizant of the way their decisions around the social framing of technical features create value tensions in a virtual space (e.g., safety vs. surveillance). Second, designers should consider adjusting the blend of technical and social features to place more responsibility for safety on direct stakeholders, namely children and their adult caregivers and less on the site itself. Third, site designers should increase the transparency of technical features and surveillance practices to facilitate user education. These implications are elaborated upon below.

Features of children’s worlds are affected by their framing, specifically by the language and practices that are encouraged by the interaction design. In Woogi World, for example, there might not be a self-organized “WWPolice” of users seeking to identify “all the bad Woogis” if there was not implicit support for this activity. This support comes in the form of 1) framing the Internet as a potentially harmful place that requires user vigilance; 2) rewarding peer monitoring by making this part of the user’s advancement structure; 3) setting apart the users who purchase memberships, creating a stratified user community (those who “make Woogi World even better” in contrast to non-paying users). While the technical feature—the ability to report violations of the rules to a site monitor—is common in children’s virtual worlds, the affect it has on user behavior may be different when it is not supported by these behavioral incentives.

In this light, separating technical from social features in discussions of children’s technology may be impossible. The argument of “safety by design” is to some extent one which employs a technological determinist framing. Safety cannot be ensured through a set of predetermined features. Instead, safety is realized through a combination of technical and behavioral interactions. Virtual worlds in their current instantiation emphasize one vision of a safe world for children, namely where they are free from “inappropriate” language. Designers should be cognizant of the way this design decision (highly constrained chat

systems) creates value tensions: while it may appeal to parent stakeholders, it comes largely at the expense of younger users’ communicative rights. It also has systemic effects: kids develop chat hacks and work-arounds, which are addressed through increased monitoring and the imposition of user penalties.

Reports from researchers (Livingstone, 2009) and government agencies (Byron, 2010; Federal Trade Commission, 2010) agree that a combination of technology, user education, and parental oversight is necessary to fulfill societal expectations of safe online play. Our analysis, however, reveals that there are different blends of these factors in children’s virtual worlds. In many virtual worlds, the technology shoulders most the safety burden, passing some responsibility to child users through peer monitoring, and almost none to the adult/parent stakeholder. Would most parents feel comfortable dropping off a preteen at an amusement park with only the park staff for oversight? The designers of children’s online worlds should experiment with the blend of technical features and stakeholder responsibility, perhaps in virtual “sandboxes”, where the level and ownership of safety features can be shifted and the effects analyzed.

Virtual worlds for children are easy to join and use, with a shallow learning curve designed to match the developmental capacity of young people. They are also opaque in the way they function, providing little guidance for users in terms of how the technical features protect their interests and keep them “safe”. For example, while all the systems we observed have limited (and evolving) chat dictionaries, users are not able to see what words are permitted or why certain, seemingly innocuous terms, are restricted. As a result of this lack of structure, the boundaries of appropriate behavior are constantly (re)discovered and contested by users. Designers should consider a balance of child agency with opportunities for teachable moments and instruction in what it means to be a good “cyber-citizen”.

Limitations and Future Research

The work presented here relies upon observational data from the investigators using the sites themselves and documenting features of the sites that they encountered during those interactions. Thus, the primary limitation of the work is that key voices are missing. There is much to be learned from site designers, children who use the sites, and their parents, educators, and other caregivers of these children. The findings from the conceptual and technical investigations presented in this work, raise pointed questions concerning the

systemic nature of virtual world interactions. Although the current work provides recommendations, these recommendations need to be bolstered by further empirical, longitudinal research. Two key questions that should be addressed through this longitudinal work with parents, caregivers, children and designers include:

- What are the implications for children’s development of self, refashionable identity within these environments?
- Can we design sites that encourage and support active parental and caregiver mediation?

Conclusion

Through this values-oriented examination of virtual worlds for children we have taken the first steps toward addressing the design tensions present in “safe” virtual spaces. While both of these worlds are “opt-in” environments – children choose to engage with these spaces – the approaches to safety and communication severely limit the creative potential within these spaces. If the hopes regarding children’s virtual spaces are to be fulfilled, the creative work will more likely be done outside these playgrounds (in affinity spaces --blogs, wikis -- and in physical spaces). Also of concern is the message these spaces may send to children about communication and interaction online: that their actions and messages are under constant surveillance by other users and anonymous, hidden site moderators. Is this an appropriate balance between parental anxiety over safety and children’s privacy? It is unclear how children perceive these safety features in use, and to what extent these features may have a “chilling effect” on interaction and communication. Additional work is planned to fill this gap in our understanding.

There is tremendous opportunity for future research in virtual worlds from a design perspective. Little is known about the lifecycle of virtual world adoption among young people, in part because these spaces are relatively new and researchers are just beginning to grasp the importance of understanding how children and virtual worlds interact. Marketers who track virtual worlds feel that these spaces are a “training ground” for next generation 3D web interfaces. As such, the social and philosophical implications should be addressed so that the systemic interactions are understood. We need to move beyond descriptions of these spaces to developing an understanding of the long-term implications of design decisions that affect children’s understanding of human values, including privacy, safety, and

communication rights. Virtual play may have real consequences.

Bibliography

- Andrejevic, M. (2007). *iSpy: Surveillance and power in the interactive era*. Lawrence, KS: University Press of Kansas.
- Beals, L., & Bers, M. U. (2009). A developmental lens for designing virtual worlds for children and youth. *International Journal of Learning and Media*, 1 (1): 51-65.
- Benderoff, E. (2007, 8 March). Cheating a real problem in Club Penguin's virtual world. *The Chicago Tribune*. Retrieved 10 April, 2008 from: <http://www.chicagotribune.com/business/chi-0703080167mar08,0,4256114.story>
- Borning, A., Friedman, B., Davis, J., & Lin, P. (2005). Informing public deliberation: value sensitive design of indicators for a large-scale urban simulation. In *Proceedings of the ninth conference on European Conference on Computer Supported Cooperative Work* (pp. 449-468). Paris, France: Springer-Verlag New York, Inc.
- Byron, T. (2010). *Safer children in a digital world: The report of the Byron Review*. Retrieved 20 April 2010, from http://www.dcsf.gov.uk/ukccis/news_detail.cfm?newsid=43&thisnews=2
- Christensen, P., & Mikkelsen, M. R. (2008). Jumping off and being careful: children's strategies of risk management in everyday life. *Sociology of Health & Illness*, 30(1), 112-130.
- Clare, E. M. (2008). Woogi World Summary for Parents. Accessed December 5, 2009 from: <http://www.WoogiWorld.com/parents/index.php?articleid=1>
- Club Penguin (2009). Parent Page. Accessed 2 December 2009 from: http://www.clubpenguin.com/parents/player_safety.htm
- Donovan, G. T., & Katz, C. (2009). "Cookie Monsters: Seeing young people's hacking as creative practice." *Children, Youth and Environments*, 19(1), 197-222. Retrieved January 19, 2010 from <http://www.colorado.edu/journals/cye>.
- Economist (23 July, 2009). Online playgrounds. *The Economist*, Retrieved 25 July 2009 from: http://www.economist.com/node/14098380?story_id=14098380
- Federal Trade Commission (2010). *Virtual worlds and kids: Mapping the risks*. [Report to Congress of the United States.] Retrieved 17 April 2010 from: <http://www.ftc.gov/os/2009/12/oecd-vwrpt.pdf>
- Flanagan, C. (2007, July). Babes in the woods. *Atlantic Monthly*, 300(1), 116-133. Retrieved

27 July, 2007, from Academic Search Complete database.

- Flanagan, M., Howe, D. C., & Nissenbaum, H. (2005). Values at play: Design tradeoffs in socially-oriented game design. *Proceedings of the conference on Human factors in computing systems (CHI '05)*. April 2-7, 2005, Portland, Oregon. New York: ACM Press.
- Flanagan, M., Belman, J., Nissenbaum, H., & Diamond, J. (2007). A method for discovering values in digital games. *Proceedings of the Digital Games Research Association Conference (DiGRA '07)* September 24-28, 2007, Tokyo, Japan.
- Friedman, B. (2004). Value Sensitive Design. In *Berkshire encyclopedia of human-computer interaction* (pp. 769-774). Great Barrington, MA: Berkshire Publishing Group.
- Friedman, B., & Kahn, P. (2003). Human values, ethics, and design. In *The Human-Computer Interaction Handbook* (pp. 1177- 1201). Mahwah, NJ: Erlbaum.
- Friedman, B., Kahn, P. H., Jr., & Borning, A. (2006). Value Sensitive Design and information systems. In P. Zhang & D. Galetta (Eds.), *Human-computer interaction and management information systems: Foundations* (Vol. 1, pp. 348-372): M.E. Sharpe.
- Friedman, B., & Nathan, L. P. (2010). Multi-lifespan information system design: a research initiative for the hci community. In *Proceedings of the conference on Human factors in computing systems (CHI '10)*, pp. 2243-2246. Atlanta, Georgia, USA: ACM. doi:10.1145/1753326.1753665
- Friedman, B., Smith, I., Kahn, P., Consolvo, S., & Selawski, J. (2006). Development of a privacy addendum for open source licenses: Value Sensitive Design in industry. In *Proceedings of the Ubiquitous Computing Conference (UbiComp '06)*. (pp. 194-211). Springer-Verlag.
- Harwood, P. G., & Asal, V. (2007). *Educating the first digital generation*. Westport, CT: Praeger.
- Ito, M., Baumer, S., Bittanti, M., boyd, d., Cody, R., Herr, B., Horst, H. A., Lange, P. G., Mahendran, D., Martinez, K., Pascoe, C. J., Perkel, D., Robinson, L., Sims, C., & Tripp, L. (2008). *Hanging out, messing around, and geeking out: Kids living and learning with new media* (John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning). Cambridge, MA: MIT Press.
- KZero Research (2009). The virtual worlds radar. Available at:

- <http://www.kzero.co.uk/radar.php>
- Linden Research (2009). Privacy policy | Second Life. Available at: <http://secondlife.com/corporate/privacy.php?lang=en>
- Livingstone, S (2009). Risk and harm. In *Children and the Internet: Great expectations, challenging realities* (chap. 6). Cambridge, UK: Polity.
- Livingstone, S., & Brake, D. (2010). On the rapid rise of social networking sites: New findings and policy implications. *Children & Society*, 24(1): 75–83.
- Lyon, D. (2001). *Surveillance society: Monitoring everyday life*. Issues in Society Series. Buckingham, UK: Open University Press.
- Marsh, J. (2010). Young children's play in online virtual worlds. *Journal of Early Childhood Research*, 8 (1): 23-39.
- Merchant, G. (2009). Literacy in virtual worlds. *Journal of Research in Reading*, 32(1), 38-56.
- Meyers, E. M. (2009). Tip of the iceberg: Meaning, identity, and literacy in preteen virtual worlds. *Journal of Education for Library and Information Science*, 50 (4): 232-242.
- Miller, J., Friedman, B., Jancke, G., & Gill, B. (2007). Value tensions in design: The value sensitive design, development, and appropriation of a corporation's groupware system. *Proceedings of GROUP Conference 2007 (GROUP '07)* (pp. 281-290). New York: ACM Press.
- Nathan, L. P. (2009). *Ecovillages, sustainability, and information tools: An ethnography of values, adaptation, and tension*. Unpublished doctoral dissertation. The University of Washington.
- Nathan, L. P., Friedman, B., Klasnja, P. V., Kane, S. K., & Miller, J. K. (2008). Envisioning systemic effects of interactive technical systems. *Proceedings of the Conference on Designing Interactive Systems (DIS '08)*, (pp. 1-10). Cape Town, South Africa. New York: ACM Press.
- Oxford English Dictionary (2009). Oxford University Press. Available at: <http://www.oed.com/>
- Palfrey, J., & Gasser, U. (2008). *Born digital: Understanding the first generation of digital natives*. New York: Basic Books.
- Pearce, C. (2009). *Communities of play: Emergent cultures in multiplayer games and virtual worlds*. Cambridge, MA: MIT Press.
- Slatalla, M. (2007, May 3). My daughter, the burger-flipping penguin. *The New York Times*.

- Retrieved 4 May, 2007, from
<http://www.nytimes.com/2007/05/03/fashion/03Online.html>
- Staksrud, E., & Livingstone, S. (2009). Children and online risk. *Information, Communication & Society*, 12 (3), 364-387.
- Thomas, D., & Brown, J. S. (2009). Why virtual worlds matter. *International Journal of Learning and Media*, 1 (1): 37-49.
- Virtual World News (21 May, 2009). "Teen and preteen migration to virtual worlds on the rise.". Retrieved 28 Nov 2009 from:
<http://www.virtualworldsnews.com/2009/05/teen-preteen-migration-to-virtual-worlds-on-the-rise.html>
- Winokur, M. (2003). The ambiguous panopticon: Foucault and the codes of cyberspace. *CTheory* (a124). Retrieved 17 April 2010 from:
<http://www.ctheory.net/articles.aspx?id=371>