MINIMUM AGE AND DEVELOPMENTAL CAPABILITIES INDICATING

READINESS FOR SELF-ADMINISTRATION WITH HOME HEALTH CARE

DEVICES OF VARYING COMPLEXITIES: AN EMPIRICAL

INVESTIGATION AND QUANTITATIVE COMPLEXITY MODEL

A thesis submitted by

Stephanie Bassock

in partial fulfillment of the requirements for the degree of

Master of Science

in

Human Factors

Tufts University

August 2016

Adviser: Daniel J. Hannon

i

Abstract

As children develop, physically and cognitively, many of them strive for personal independence. For children diagnosed with chronic diseases, their desire to act independently might lead them to become involved in their own healthcare, performing diagnostic and therapeutic tasks using medical devices. This research identified a typical age at which children are more capable of performing tasks with home health care devices (i.e., blood glucose meters, nebulizers) with minimal error. In addition, this research evaluated the relation of child development and device complexity to the prevalence of use errors. Nine years of age emerged as a threshold at which a majority of children could perform medical device tasks with minimal error. Moreover, children's age and working memory capacity and device complexity accounted for a significant proportion of the variance in use error rate. Additionally, the researcher provided a method for quantifying device complexity as well as a metric for estimating the rate of potential for use errors by children.

Acknowledgments

Table of Contents

Table of Figures

Table of Tables

MINIMUM AGE AND DEVELOPMENTAL CAPABILITIES

INDICATING READINESS FOR SELF-ADMINISTRATION WITH HOME

HEALTH CARE DEVICES OF VARYING COMPLEXITIES: AN EMPIRICAL

INVESTIGATION AND QUANTITATIVE COMPLEXITY MODEL

1. Introduction

As children develop, physically and cognitively, many of them strive for personal independence. Children develop a sense of industry as they meet developmental milestones and accomplish new tasks, such as assuming academic responsibilities, regulating emotions, and interacting with peers appropriately (Erikson, 1956; Howe et al., 2011). For children diagnosed with chronic diseases, their desire to act independently might lead them to become involved in their own healthcare. For example, children might request to perform self-monitoring tasks or administer their medication independently. However, children's developmental capabilities play a critical role in the success of their independence. To assure the child's safety, it is imperative that parents and clinicians have a comprehensive understanding of the child's biological and cognitive capabilities before relinquishing self-care tasks (Clarke, 2011; Scott, 2013). Releasing health-related responsibilities prematurely can lead to poor disease management (McNally, Rohan, Pendley, Delamater, & Drotar, 2010). However, there is limited literature available on the age and developmental variables that predict readiness. Such knowledge would enhance parent's and clinician's assessments of whether children are developmentally equipped to assume health-related tasks that are central to their wellbeing; tasks that need to be performed correctly or else harm could occur.

A better understanding of the developmental prerequisites to readiness would enable medical device manufacturers to design devices that are more

developmentally appropriate for children. Two of the most common chronic diseases among children – type 1 diabetes and asthma – require the use of home health care devices to manage symptoms effectively. Specifically, diabetes requires frequent blood glucose monitoring with a blood glucose meter and multiple daily insulin dosing with a vial and syringe, insulin pen, or pump. Asthma typically requires the use of one or more inhalation devices, such as a metered dose inhaler or nebulizer. Unfortunately, many of these medical devices do not account for the variability in developmental skill mastery among children. Rather, they appear to be designed for use by adults even though they might be indicated for adults as well as children.

Although developmental changes are relatively predictable, each child has a unique trend of growth with an individual personality, learning style, and experiential background (Bredekamp, 1987). Therefore, certain medical device elements might create a barrier to performing autonomously. For instance, the medical device might contain components that require strength exceeding that of the child or text that is above the child's reading level. A better understanding of the developmental milestones that predict readiness will inform the development of device user interfaces and ultimately, ensure that children who are ready to manage aspects of their own care have the resources to do so with minimal chance of making an error.

The notion of accounting for user characteristics when designing medical devices echoes the U.S. Food and Drug Administration's (FDA) philosophy.

However, it merely speaks to one major element the FDA recommends medical device manufacturers consider when developing devices. To inform manufacturers of the practices that promote safe and effective medical device use, the FDA document entitled, "Applying Human Factors and Usability Engineering to Medical Devices" (U.S. Food and Drug Administration, 2016) identifies three major human factors engineering considerations for error prevention. Medical device manufacturers should account for characteristics subject to the devices' intended users, use environments, and device user interface. This thesis focuses on the user- and device-specific variables that might pose a barrier to proper medical device use by children.

In addition to evaluating the extent to which children's developmental capabilities affect performance, the researcher assessed how a characteristic subject to the device user interface, the device complexity level, affects use error rate. Variability in user interface complexity might influence children's ability to use medical devices effectively. There is a wide array of medical devices representing varying levels of complexity in today's market. The devices range from simple (e.g., thermometers, peak flow meters) to complex (e.g., infusion pumps, ventilators). Several studies have shown that user performance decreases as task complexity increases (Just & Carpenter, 1992; Anderson & Jeffries, 1985; Sohn & Doane, 2003). The result is likely applicable to device complexity, noting that highly complex tasks and device both comprise a relatively high number of information elements (e.g., functions, information

cues). Decreases in performance were attributed to the complex tasks' excessive demand on working memory (i.e., short-term memory). The tasks contained a high number of information elements and subsequently, required a cognitive capacity exceeding that of the participants. These findings suggest that medical devices with higher complexity are potentially more cognitively demanding and, as a result, increase the chance of errors. Thus, the researcher provides a method for quantifying device complexity to enable the assessment of the relationship between medical device complexity and use error rate.

In short, the researcher's first aim was to examine the relationship between child development and children's readiness for managing aspects of their health care. Specifically, it would achieve this by (1) identifying the transition point in development (i.e., chronological age) that indicates a child is capable of using home health care devices (i.e., blood glucose meters, nebulizers) with a minimal chance of making an error and (2) evaluating the extent to which child development variables contribute to use error rate. The second aim was to develop a method to quantify a device's level of complexity and relate device complexity to the prevalence of use errors by children.

2. Survey of Literature on Child Development

The Association for the Advancement of Medical Instrumentation (AAMI; 2009) HE75 defines a use error as an "undesirable or unexpected event resulting from the interaction between a user and a device." The researcher utilizes use error rate to gauge the level of risk associated with using home health care devices, and ultimately estimate children's readiness for successful chronic disease management.

Asthma and type 1 diabetes are particularly prevalent chronic diseases among children. As of 2014, approximately 6.3 million children in the United States had been diagnosed with asthma (Centers for Disease Control and Prevention, 2014). Moreover, about 84,100 U.S. children received a type 1 diabetes diagnosis (International Diabetes Federation, 2015). As such, children commonly rely on home health care devices to manage their symptoms, such as blood glucose meters, insulin pumps, insulin pens, nebulizers, and inhalers. The following subsections expand on biological and cognitive aspects of child development that might influence use error rate with the abovementioned home health care devices. Specifically, the subsections address the following developmental factors:

- Biological aspects
  - o Neurological development
  - o Hand function
  - o Hand size

- Cognitive aspects
  - Understanding of numerosity
  - Language development
  - Working memory capacity

In addition, this section discusses the literature on the relationship between age and children's readiness to self-administer treatment with home health care devices.

2.1. Biological Development

Biological processes refer to changes in a child's body, such as the development of neurological and motor function, as well as changes in body size (Santrock, 2013, p.13). The literature on neurological development suggests that children as young as six years of age are capable of participating in aspects of their own care.

2.1.1. Neurological Development

Lenroot and Giedd (2006) utilized anatomical magnetic resonance imaging to study brain development in children (as cited in Santrock, 2011). By age six, a child's brain has grown to approximately 95% of its adult size. The most rapid growth was shown to take place in the frontal lobe, which contributes to the development of executive function. Such functioning impacts children's ability to sustain attention in demanding tasks as well as plan and organize new actions (Gogtay & Thompson, 2010; Munakata, Casey, & Diamond, 2004).

Previous research has shown that children's executive functioning level is associated with adhering to diabetes treatment routines and their ability to self-manage their disease (McNally, Rohan, Pendley, Delamater, & Drotar, 2010). Children with relatively high executive functioning were more likely to adhere to treatment and ultimately hit more glycemic targets.

Furthermore, around six years of age, the temporal and parietal lobes begin to accelerate in growth. The temporal and parietal lobes attribute to the acquisition of various cognitive processes, such as the development of language and spatial abilities. For children diagnosed with a chronic disease, well-developed language skills likely facilitate children's comprehension of medical device labeling or their understanding of device training. Moreover, children's capacity to understand spatial relations likely aids in recognizing which medical device components fit together during assembly, as spatial abilities support differentiating among shapes.

However, past work on medical device use by children showed that children were unable to perform tasks correctly until roughly 11 years of age (Naughten, 1982; Perwien, Johnson, Dymtrow, & Silverstein, 2000). Perwien et al. (2000) examined children's blood glucose testing skills and found that 10.9 was the mean age of children who were able to perform glucose monitoring tasks. Tasks included obtaining a sufficient sample of blood and applying the sample to the test strip properly. Furthermore, Naughten (1982) conducted a study examining the relationship between age and children's ability to use insulin

administration devices properly. The findings were similar to the previous study in that children were able to perform tasks with insulin administration devices successfully around 11.2 years of age. Although the literature on neurological development indicate that the brain develops substantially early in life, there seem to be additional factors creating a barrier to children's mastery of self-care tasks.

2.1.2. Hand Function and Size

According to Santrock (2013, p.106), during middle and late childhood (i.e., six to 12 years of age), children's hand strength and manual dexterity develop and refine considerably. It is not until around eight to ten years of age that children develop the ability to coordinate their fingers to manipulate objects with ease and precision. Children younger than eight might experience difficulty performing medical device tasks that require well-developed hand function. For example, children might not possess the fine motor skill to remove a blood glucose meter test strip from its vial with ease, or sufficient hand strength to remove a cap.

Furthermore, children's hand size might affect their ability to perform tasks with medical devices. Children with a relatively small hand size might be at a disadvantage depending on the extent to which the device was designed to fit in an average adult hand. For example, children with a hand length shorter than that of a metered-dose inhaler's height might experience difficulty taking an asthma treatment due to their inability to reach the inhaler's canister.

Consequently, children's hand function and size might have affected their ability to perform the aforementioned glucose monitoring and insulin administration tasks successfully.

2.2. Cognitive Development

Many medical devices require children to possess complex cognitive skills. For instance, medical devices tend to require sufficient working memory capacity for effective decision making and problem solving, as well as the literacy knowledge to read and comprehend device labeling. Specifically, blood glucose monitoring requires the ability to perform numeric operations with three digit numbers to interpret the results correctly. Effective administration of asthma medication requires measurement skills for calculating doses.

2.2.1. Functional Numerosity Development

Piaget, Inhelder, and Weaver (1969) studied cognitive development in children and found that individuals who are seven to 11 years of age develop the ability to perform tasks that require concrete operations. Such operations pertain to actions that are mentally reversible, also referred to as conservation. Piaget et al. (1969) indicated that children who possess the cognitive ability to understand conservation are more likely to grasp concepts related to demonstrating functional numerosity (e.g., numeric ordering, serialization).

Conservation refers to a child's ability to form a scheme of and mentally manipulate something that is not physically present. A standard study that

demonstrates the concept of conservation involves presenting a child with two identical clay balls. The experimenter rolls one clay ball into a long, thin strip, and asks the child whether the two pieces (i.e., clay ball vs. clay strip) contain the same amount of clay. A child who has not yet acquired the ability to perform concrete operations would reply that the long strip consists of more clay than the ball. To respond correctly, children need to mentally reverse the action by envisioning the clay's size as if it were rolled back into a ball. Children do not normally develop this ability until around seven or eight years of age. Many tasks related to self-managing chronic disease symptoms require children to perform concrete operations, such as adding or comparing number values. For instance, a child performing a blood glucose test needs to determine whether the blood glucose value is too high or low relative to a value considered to be in the healthy range.

However, there is a potential barrier to adequate administration by children who have mastered concrete operations. Although Piaget's research on conservation suggests that children as young as seven are capable of performing the mathematical operations required to use their devices successfully, the Common Core Standards (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010) do not introduce these concepts until second grade. Therefore, children are typically eight or nine years of age when they master mathematical operations related to measuring and algebraic thinking. Consequently, children might commit numerosity-related use

errors because they have not learned mathematical concepts for which they are capable of understanding.

2.2.2. Language Development

As children develop the numerosity skills required to reason logically, they become increasingly able to communicate effectively (Santrock, 2013). For example, children who possess the ability to form concrete operations are more likely to use comparatives appropriately in speech (e.g., "My blood sugar is higher than it's supposed to be").

Another aspect of language development that supports children's ability to participate in health-related tasks concerns literacy level. Early in elementary school, children are taught the fundamentals of reading, such as the development of rudimentary vocabulary and grammar concepts. However, the taxing demand reading imposes on emergent readers results in minimal resources available to comprehend the material.

Once children reach fourth grade (nine to ten years of age) they have typically acquired the reading skills necessary to fully understand printed material. However, the extent to which the child comprehends the text is dependent upon its appropriateness given the child's reading level. For example, children who read at a fourth-grade level are less likely to understand a medical device's instructions-for-use (IFU) written at a seventh-grade reading level. Unfortunately, medical device labeling might create a barrier to proper medical device use considering that the FDA document entitled, "Guidance on Medical

Device Patient Labeling," (U.S. Food and Drug Administration, 2001) recommends medical device manufacturers design the device labeling such that the text is written at an eighth-grade level or below. Consequently, the available learning aids might exceed a child's reading level.

2.2.3. Working Memory Development

Working memory is responsible for storing and manipulating information in short-term memory (Myatchin & Lagae, 2013). According to Miller (1956), working memory has a limited capacity. Humans can store and process seven, plus or minus two pieces of information. Cowan (2004) conducted an extensive review of literature succeeding Miller's (1956) research on working memory capacity and indicated that an average adult has a capacity of about four items, and fewer in children. Children's working memory capacity is comparable to that of adults, approximately three to four items, at ten years of age (Riggs, McTaggart, Simpson, & Freeman, 2006). Several studies identified a relationship between working memory capacity and children's performance attending to and carrying out complex cognitive tasks and daily activities. For example, developmental increases in working memory have a positive effect on children's ability to read, solve problems, and perform mental calculations (Bull & Scerif, 2001; Hitch, Towse, & Hutton, 2001). Therefore, working memory capacity might contribute to children's ability perform health-related tasks. For example, children with a working memory capacity greater than three items might be more apt to comprehend medical device learning aids (e.g., IFU, on-screen

prompts, training), measure medication doses accurately, and perform device steps in the correct sequential order (e.g., remembering to wash hands prior to lancing for glucose testing).

2.3. Summary of Child Development Variables

The literature on child development identified that there is inevitable variability among children of the same chronological age (Sroufe, Cooper, DeHart, & Bronfenbrenner, 1992). Thus, it is unlikely that all children with chronic diseases converge in their ability to self-manage their illness at the same age. Individual children develop biologically and cognitively at different rates. Therefore, one or more developmental milestones might influence readiness.

Subsequently, the researcher sought to identify the developmental variables that tend to plateau during childhood. The National Institutes of Health (NIH) collected data using the NIH Toolbox for the Assessment of Neurological and Behavioral Function (Hodes, Insel, Landis, & On behalf of the NIH Blueprint for Neuroscience Research, 2013) and found that data on children's dexterity and working memory capacity supported non-linear relationships with age (Tulsky et al., 2013; Wang, Bohannon, Kapellusch, Garg, & Gershon, 2015). As such, children with more developed fine motor skills and greater working memory capacities might commit relatively fewer use errors. In addition, the literature led to the consideration of grip strength, hand size, reading ability, and functional numerosity as variables that might influence use errors rates.

3. School Nurse Interviews

A critical role of a school nurse is to provide health care to children with chronic diseases. Therefore, the researcher completed interviews with two school nurses to inform the primary study's design. The main objectives for conducting the interviews were to (1) identify the types of home health care devices and chronic diseases school nurses encounter most frequently, and (2) gain a clinical perspective on children's readiness to manage aspects of their own care.

The most common chronic diseases the school nurses encounter are type 1 diabetes and asthma. The school nurses reportedly provide support to most of those children on a daily basis and utilize several types of devices to aid in symptom management, including blood glucose meters, insulin pumps, inhalers, nebulizers, and peak flow meters. Therefore, the researcher determined that the primary study include devices intended for use by children with diabetes or asthma to increase the generalizability of the results.

Furthermore, both school nurses expressed concern regarding the lack of training children receive on their devices. In their experience, children receive little to no training on their device unless a parent or caregiver opts to bring the child to a respiratory or diabetes clinic. One nurse opined that the lack of training results in a clear increase in error. For example, she has observed several children with diabetes develop infections because they did not wash their hands before lancing themselves for glucose testing.

The researcher asked the school nurses to explain how they assess whether a child is capable of participating in aspects of their own care. Both school nurses responded that the strongest indications of success are children's cognitive and emotional abilities. Children's reading ability and understanding of numerosity appear to pose as barriers to performing autonomously with home health care devices. One nurse explained that many children with asthma are unable to read the labeling on albuterol packaging. Accordingly, some children fail to differentiate among the packages and tend to bring the wrong medication type to school. Moreover, some children experience difficulty reading the inhaler dose counter's value. The children are unable to determine whether there is medication left in the inhaler and, as a result, fail to alert their parent or caregiver that they need a new canister.

In addition, the school nurses explained that children's ability to manage their stress and anxiety plays a role in using home health care devices successfully. However, aspects of emotional development appeared to pertain specifically to devices that contain needles, such as auto-injectors and pre-filled syringes.

Lastly, one nurse said that children's physical development influences home health care device use. She explained that many children with diabetes experience difficulty depositing enough blood onto the test strip to activate the meter due to the strip's small size and the children's limited dexterity.

The interview responses fortified the researcher's presumption that aspects of cognitive and biological development influence children's ability to use home health care device safely and effectively. As such, the researcher designed the primary study to include assessments of children's reading ability, understanding of numerosity, fine motor ability, as well as several additional aspects of development identified in the literature.

Notably, the researcher did not evaluate children's emotional intelligence during the primary study because it was outside the scope of this thesis. While children's emotional development likely impacts task performance, it is fairly difficult to assess, noting that valid measures of emotional intelligence require a great deal of time to administer.

4. Review of Literature on Design Complexity

Devices with relatively high complexity might increase use error rate, considering that a standard method for increasing the usability of a device is to minimize excess complexity (Association for the Advancement of Medical Instrumentation, 2009). To test this theory, the researcher sought to identify a method for quantifying the complexity of home health care devices. Unfortunately, a product class-specific method did not appear to exist in the literature. Therefore, the researcher surveyed the literature on complexity metrics related to a seemingly relevant concept called design complexity.

Design complexity pertains to the extent to which a given device or system will affect the amount of effort (i.e., man-weeks) required by the designer (Crespo-Varela, Medina, & Kremer, 2012). Several studies have identified methods for assessing design complexity within a variety of domains, such as software, electrical, and product (Ameri, Summers, Mocko, & Porter, 2008; Claasen, 2003; Zhang, Li, & Tan, 2010). Crespo-Varela et al. (2012) surveyed the literature on design complexity to identify the methods potentially applicable to calculating medical device design complexity. Crespo-Varela et al. (2012) included design complexity measures based on their generalizability. The aim was to identify the design complexity measure(s) most predictive of the FDA's decision time of approval. They utilized the device manufacturer's 510(k) submission date and FDA's decision date to calculate decision time of approval. Crespo-Varela et al. (2012) deemed the methods developed by Bashir and Thomson (1999) and

Roy, Evans, Low, and Williams (2011) as valid measures to quantify medical device design complexity. The design complexity scores the metrics yielded were strongly associated with the FDA's decision time of approval.

Bashir and Thomson's (1999) metric pertains to device functionality. The procedure involves decomposing device functions into hierarchical levels. The top level represents basic functions, and subsequent levels denote subfunctions. Bashir and Thomson (1999) illustrate the concept by decomposing a battery charger's functions. They indicated that the battery charger's basic, top-level function was supplying DC power. They proceeded by decomposing the basic function into subfunctions, which included assembling components, protecting the device, converting power, and providing an interface. Bashir and Thomson (1999) broke down each sub-function until they could not be decomposed further. To quantify design complexity, they implemented Equation 1.

Equation 1: Design complexity

$$Design\ Complexity = \sum_{j=1}^{l} F_j * J$$

where,
F_j is the number of functions at level j
l is the number of levels

Essentially, for each level, they multiplied the number of functions at that level by the level's position in the hierarchy. Then, they calculated the sum of the products. As such, complexity is an objective measure of device functionality,

such that increasing the number of device functions leads to an increase in complexity level.

Conversely, Roy et al. (2011) measured design complexity based on the commonality of device components. Specifically, Roy et al. (2011) quantified design complexity according to the number of device part variations and total number of combinations, which they referred to as the design ratio. Equation 2 illustrates the method for calculating the design ratio. They considered device parts with a relatively low design ratio as more complex.

Equation 2: Design ratio

$$Design\ ratio\ (i) = \frac{n_i}{n}$$

where,
n_i = number of product variants that use part variant i
n = total number of product variants

Roy et al.'s (2011) metric concerned device variants, a logical consideration when the aim is to reduce design complexity and ultimately, decrease device manufacturing cost and design time. However, the researcher was unable to find theoretical evidence that would reasonably support a hypothesis involving the design ratio as a significant predictor of use error rate. Therefore, Roy et al.'s (2011) metric is not discussed further.

Ultimately, Bashir and Thomson's (1999) metric appeared to be a promising method for predicting use error rate due to its function-driven nature. Increasing device functionality generally leads to a rise in complexity (Norman,

2002). Increasing the number of device functions likely increases the number of

information cues processed and distinct actions executed by the user. As

mentioned in a previous section, the brain is limited in its capacity for processing

and storing information. Increasing the number of information elements the user

has to attend to while performing a task decreases performance (Anderson &

Jeffries, 1985; Just & Carpenter, 1992; Miller, 1994; Sohn & Doane, 2003).

Subsequently, a device with a relatively high number of functions and

components might impose excessive cognitive demand on working memory,

leading to a decrease in task performance.

5. Device Complexity

5.1. Device-User Complexity Model

The literature identified Bashir and Thomson's (1999) functional decomposition model a valid technique for estimating the amount of time a designer will spend developing a new device. To increase the likelihood that the model is predictive of use error rate, the researcher modified the functional decomposition procedure. The new model assumes a more user-centered process. Unlike Bashir and Thomson's (1999) model, the new model omits functions that map to components with which the end user does not interact, such as a mechanical clock's motor. Figure 1 illustrates the modified functional decomposition model workflow, hereafter referred to as the Device-User Complexity (DUC) model.
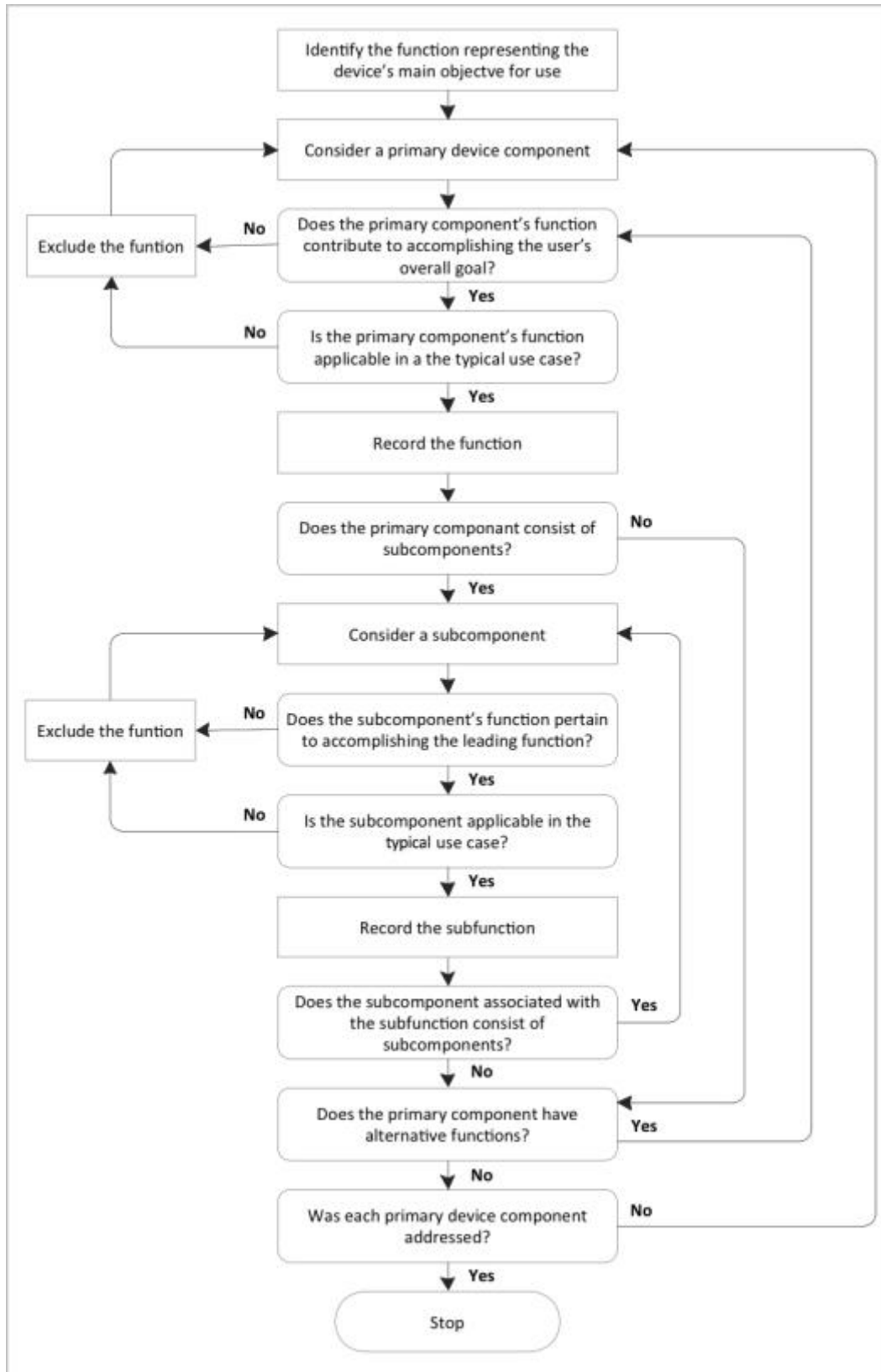
Figure 1: Device-User Complexity (DUC) workflow diagram

Figure 1 provides a higher-level visual representation of the following steps:

Step 1. Record the function that represents the device's main objective for use at the functional tree's first hierarchical level. For example, the user's overall goal for using a wrist watch is to monitor the current time continuously. Figure 2 illustrates a wrist watch Device-User Complexity (DUC) functional tree.
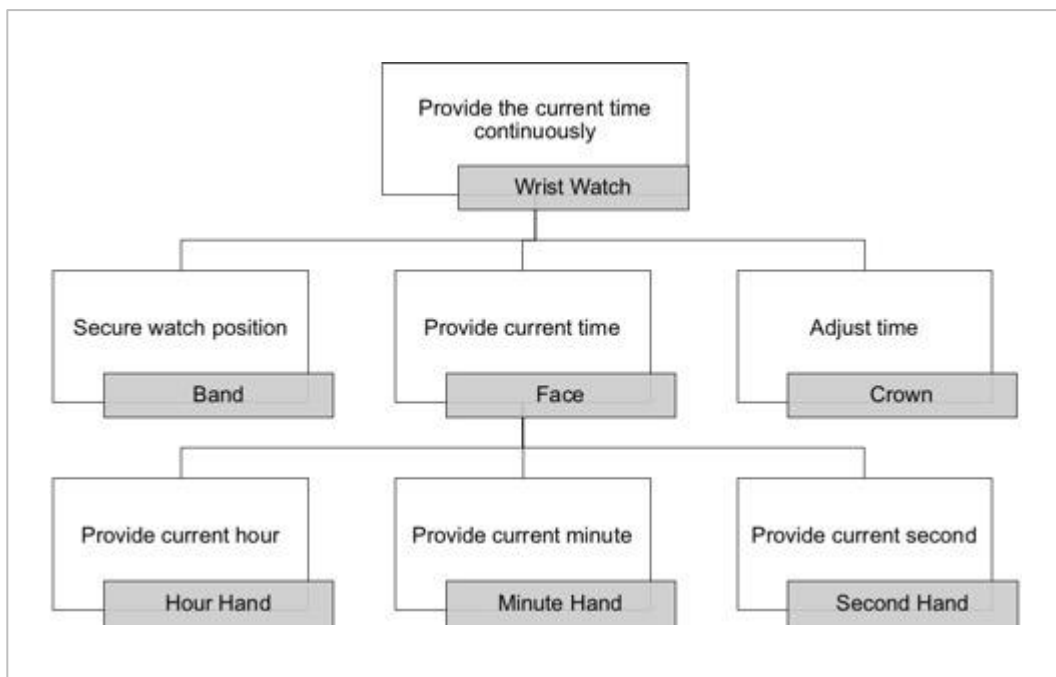
Figure 2: Wrist watch DUC functional tree

Step 2. Consider a primary device component. When considering a wrist watch, such as the one presented in Figure 3, the primary component might be the watch's wristband, face, crown, or battery cover.

Figure 3: Wrist watch primary components

Step 3. Identify whether the component's function contributes to accomplishing the user's overall goal. To increase the likelihood that use error rate is predictive of the Device-User Complexity (DUC) score, the model only includes functions that pertain to actions that are:

- Necessary for accomplishing the device's primary objective for use.

- Applicable in the most common use case.

For example, a watch's face provides the time, enabling the user to achieve the ultimate goal of monitoring the current time continuously. Moreover, a watch user interacts with the watch's face frequently. Therefore, the component meets the aforementioned requirements and included in the DUC functional tree at the second hierarchical level.

Conversely, the watch's battery cover relates to an auxiliary function. The action associated with the component's purpose – changing the battery (action) to restore power (component's purpose) – is a low-frequency maintenance action. Furthermore, the initial placement and replacement of a watch battery is not a typical task of the primary user. As such, users are less likely to commit use errors related to replacing the battery, and thus excluded from the DUC functional tree.

Notably, the DUC functional tree excludes functions associated with static device labels or markings. Typically, designers implement such elements to prevent use errors (Association for the Advancement of Medical Instrumentation, 2009). For instance, tick marks drawn on a watch's face facilitate accurate time measurements. As such, static labels and markings that are designed appropriately are presumably less likely to influence device complexity.

Step 4. Decompose the function further, when appropriate. To determine whether the function affords further decomposition, identify whether the primary component associated with the function consists of subcomponents. In Figure 4, the watch's face, a primary component, has three subcomponents, including the hour hand, minute hand, and second hand.

When there are additional subcomponents to address, proceed to Step 5. Otherwise, continue with Step 8.

Figure 4: Wrist watch subcomponents

Step 5. Consider a particular subcomponent. In Figure 4, the hour hand is a subcomponent of the watch's face.

Step 6. Determine whether the subcomponent's function pertains to accomplishing the leading function, and applicable in a typical use case. In the wrist watch example, the hour hand supplies the current hour, which enables the watch's face to provide the time of day. Moreover, a watch user interacts with the watch's hour hand frequently. Therefore, the DUC functional tree includes the hour hand's function. Place a subfunction that meets the requirements at the following hierarchical level in the DUC functional tree. Exclude a subfunction that does not fulfill the requirements.

Step 7. Decompose the subfunction further, if possible. When the subcomponent associated with the leading subfunction consists of one or more

subcomponents, repeat Steps 5 – 7. After decomposing all applicable

subfunctions, proceed with Step 8.

Step 8. Determine whether the primary component has alternative

functions. When the primary component has one or more unaddressed

alternative functions, repeat Steps 3 – 8 for each. Then, continue to Step 9.

A single component can have several functions. Consider a digital watch

that displays the current time, a stopwatch, and timer. The watch has a button

that enables the user to toggle among the three display modes. Thus, the single

component has three functions, each of which afford decomposing. According to

Weinger, Gardner-Bonneau, Wiklund, & Kelly (2011) "…with multifunctionality

comes increased cost and complexity, decreased usability, and increased risk of

use errors" (p. 700). Therefore, the DUC score accounts for all applicable device

functions.

Step 9. Repeat Steps 2 – 8 until the DUC functional tree addresses all

primary device components. Then, compute the DUC score by employing Bashir

and Thompson's (1999) equation. Figure 5 and Equation 3 demonstrate the

procedure.
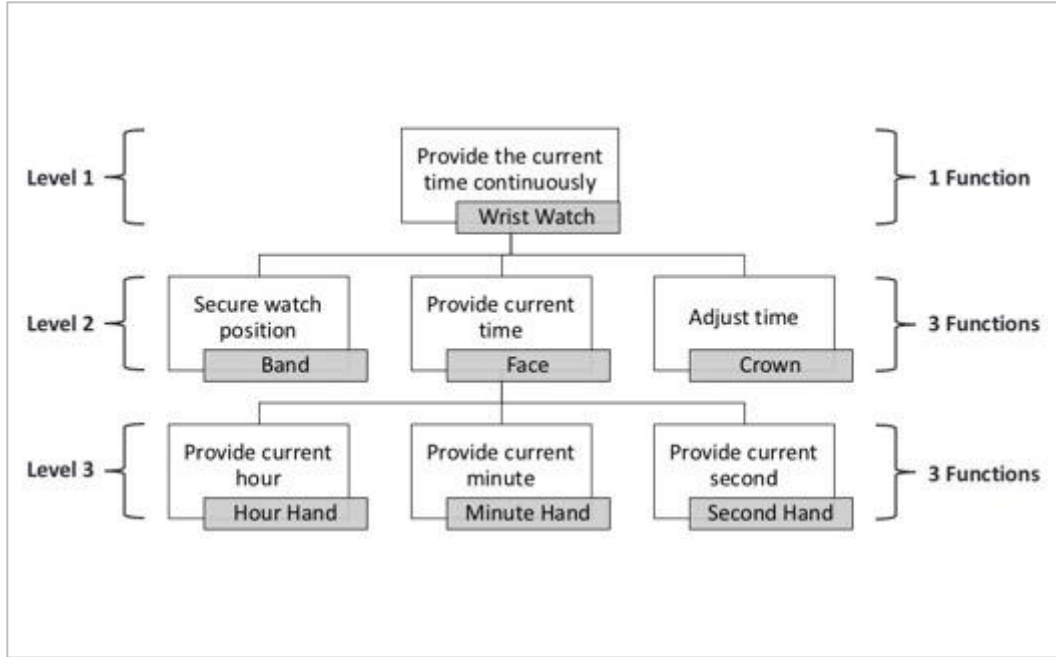
Figure 5: Wrist watch DUC functional tree

Equation 3: Wrist watch DUC calculation

$$DUC = \sum_{j=1}^{l} F_j * J$$

$$\boldsymbol{DUC = (1)(1) + (3)(2) + (3)(3) = 1 + 6 + 9 = 16}$$

For each level, multiply the level's hierarchical position by the number of functions at that level. In Figure 5, there is one function at the first hierarchical level, (1)(1) = 1, three functions at the second hierarchical level, (3)(2) = 6, and three functions at the third hierarchical level, (3)(3) = 9. The DUC score is the summation of the products, 1 + 6 + 9 = 16. The score represents a given device's complexity level. The researcher conducted a preliminary study to valid the DUC model.

5.2. Preliminary Study

5.2.1. Purpose

The purpose of conducting the preliminary study was to validate the DUC model for Human Factors application. The metric was adapted from Bashir and Thomson's (1999) functional decomposition method, which estimates the number of hours required to design a product. The main objective of developing the DUC model was to determine the extent to which device complexity level predicts the potential for use errors. The researcher sought to assess the DUC model's validity by examining the relationship between participant perceived complexity and device complexity scores, which the researcher quantified using the DUC procedure.

5.2.2. Method

5.2.2.1. Participants

The researcher recruited twelve participants from Tufts University Human Factors classes. Participants received course credit for participating in the study. The researcher assigned all participants a numerical identifier to preserve anonymity. The average age of participants was 22.33 years (SD = 1.55). Furthermore, the sample of participants included seven females and five males.

5.2.2.2. Devices

Participants evaluated six devices from three device categories, including: two kitchen thermometers (Figure 6 and Figure 7), watches (Figure 8 and Figure 9), and coffee makers (Figure 10 and Figure 11). The researcher calculated the DUC score for each device (Appendix A.). Each device had a distinct score.

Figure 6: Taylor® Classic Instant Read Pocket Thermometer (Kitchen Thermometer 1) – DUC 8

Figure 7: Oneida® Digital Probe Thermometer (Kitchen Thermometer 2) – DUC 44

Figure 8: Kate Spade® Watch (Watch 1) – DUC 14

Figure 9: Casio® Sport Watch (Watch 2) – DUC 66

Figure 10: Mr. Coffee® TF5 Coffee Maker (Coffee Maker 1) – DUC 29

Figure 11: Cuisinart® Grind & Brew Automatic Coffee Maker (Coffee Maker 2) – DUC 112

5.2.2.3. Experimental Protocol

All experiments took place in the Human Factors Usability Laboratory Observation Room at Tufts University (Figure 12). Upon arrival, the researcher provided participants with the informed consent form. Participants were given sufficient time to review the form before consenting to participate in the study.

Figure 12: Human Factors Usability Lab Observation Room

The researcher presented participants with six devices as well as the

devices' associated IFUs. Next to the devices was a Windows 7 desktop computer

that displayed a Qualtrics survey. The researcher instructed participants to assess

the complexity of each device and then respond to survey questions. To ensure

participants understood the study procedure, the researcher walked each

participant through three practice questions before allowing the participant to

work independently. Notably, the researcher instructed the participants to use

the IFUs solely as a resource for developing an understanding of each device's

interface, rather than a factor in their assessments.

Each IFU was different regarding its legibility and layout. To ensure the

IFUs' designs did not bias the participants' responses (e.g., rating a device as

more complex due to the IFU's poor legibility), the researcher revised the IFUs

such that the layouts were presented in a consistent format. To increase

participants' understanding of how each device functioned, each revised IFU

displayed three to four images of the device, a list of device components, and

step-by-step manufacturer instructions on how to use the device.

5.2.3. Survey

The survey questions asked participants to assess the complexity of each

device. Specifically, participants were shown two devices at a time and prompted

to compare the devices' complexity levels on a sliding scale. Each survey question

displayed one device on either side of the scale (Figure 13).

Adjust the slider to indicate how much more complex one item is than the other. If they have the same complexity, leave the slider in the middle.

Thermometer 2   Coffee Maker 2

Thermometer 2 is much more complex   Equal complexity   Coffee Maker 2 is much more complex

Figure 13: Survey sample question

Adjusting the slider's bar toward a device signified the extent to which

the participant perceived that device as more complex than the other. For

example, pushing the slider bar slightly to the right indicated that the device on the right was slightly more complex than the device on the left. Notably, leaving the slider in the scale's center indicated that the participant perceived the two devices as equally complex.

The researcher designed the survey to emulate a tournament. All participants responded to all possible device pairings. For example, participants compared the digital thermometer to the five other devices. Notably, the survey presented the device pairings in random order to prevent order effects.

Participants compared devices on a continuous scale from negative five to five (Figure 14). The actual survey did not display the value labels in Figure 14. The researcher allocated or "awarded" the quantity to the slider bar's left to the device on the right and vice versa (Figure 15).
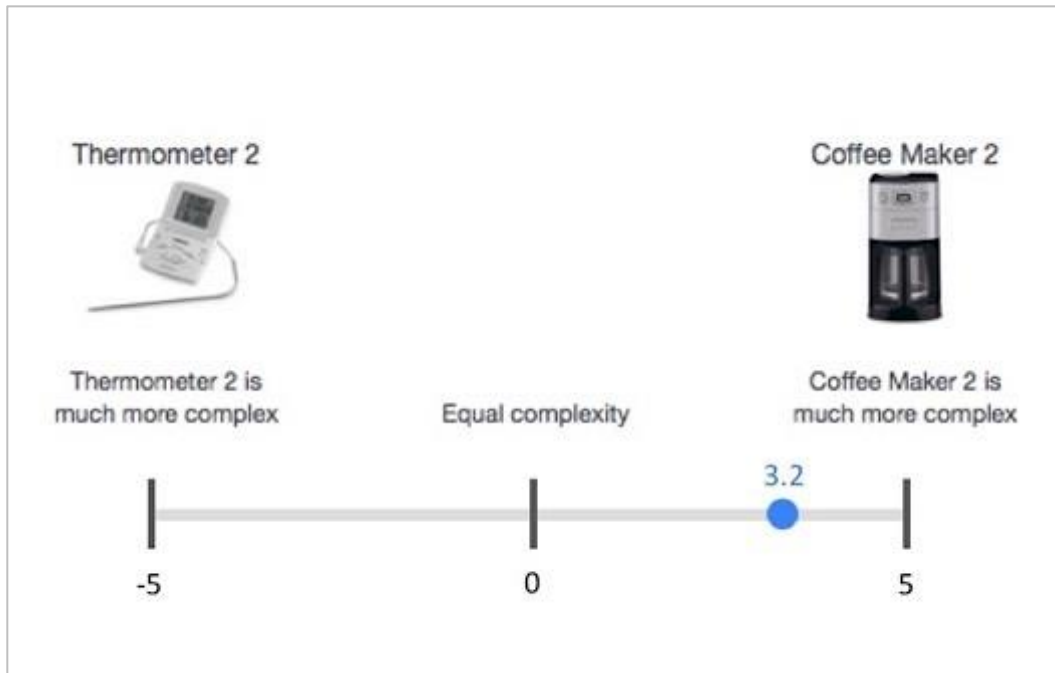
Figure 14: Survey measurement scale
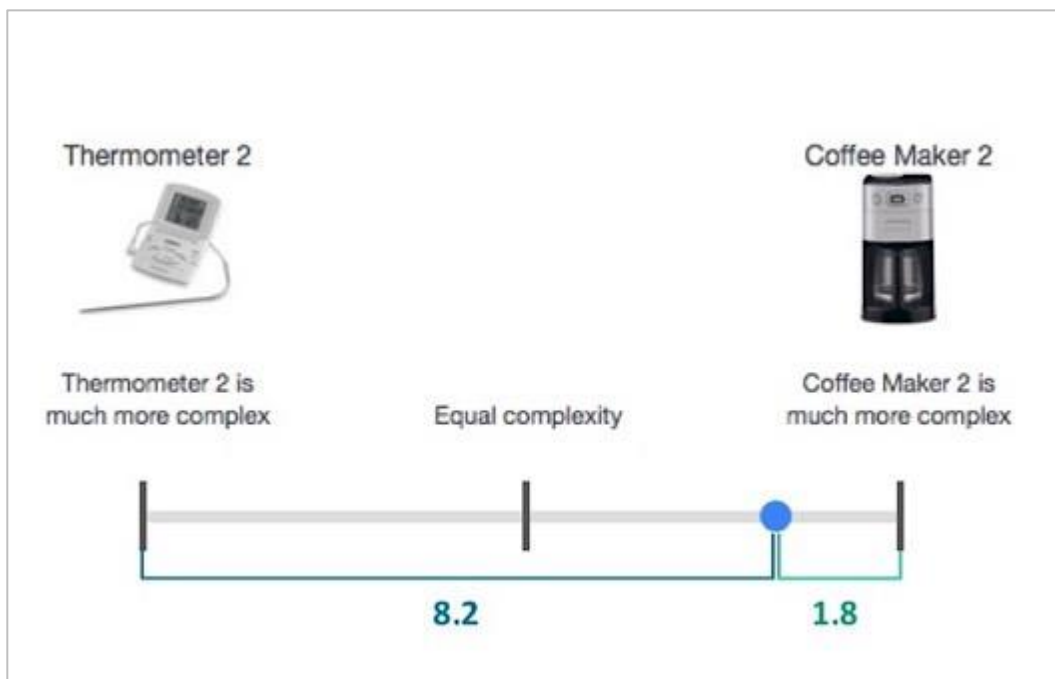
Figure 15: Survey measurement scale score allocation

Each device appeared five times in the survey. To calculate a final device

complexity rating, the researcher summed participants' five ratings of a

particular device. Therefore, participants' final device complexity ratings of the

six devices produced a ranking. Importantly, the ordinal values possessed continuous variable properties. A typical ranking scale merely denotes the one-point difference between any two items. Eliciting head-to-head comparisons on a continuous measurement scale was advantageous because the final complexity ratings indicated the extent to which participants perceived a device as more complex than another, and not purely a rank difference.

5.2.4. Results and Discussion

The tournament process for head-to-head comparisons produced a distribution of ratings that reflected the perceived magnitude of the difference between any device and all other devices in the tournament. Figure 16 provides the average final complexity rating for each device.
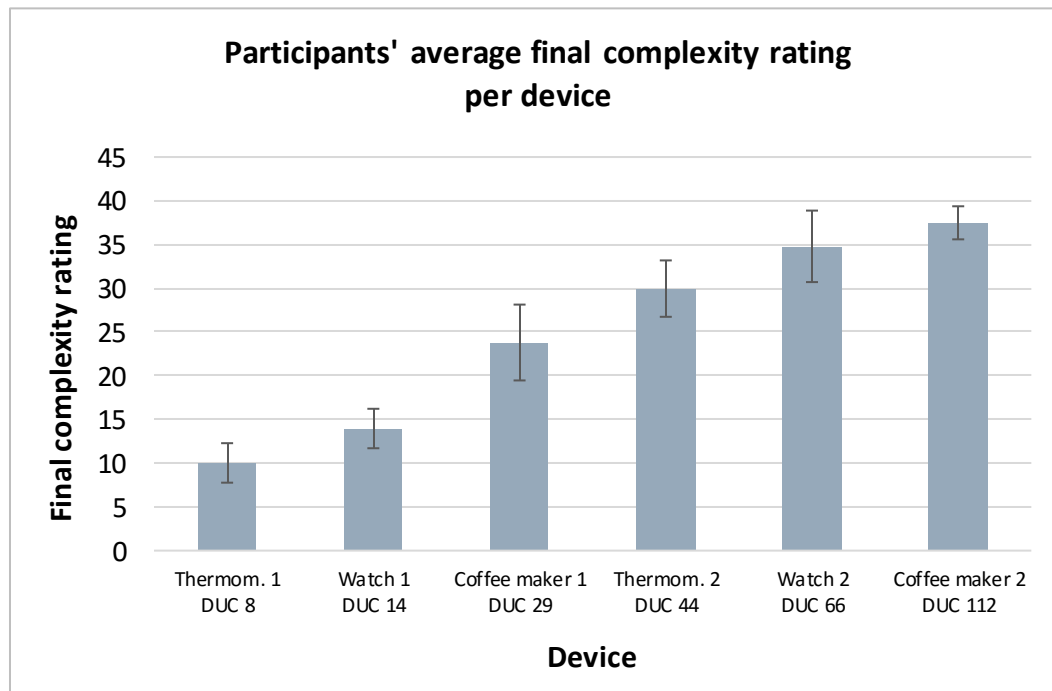
Figure 16: Participants' average final complexity rating per device. Bars indicate +/- 1 standard deviation.

The reasonably low standard deviation for each device suggests that there was high agreement among participants regarding their assessments of device complexity. To determine the degree of participant agreement, the researcher evaluated the interrater reliability using intraclass correlation coefficients (ICC) (McGraw & Wong, 1996). The ICC for participant complexity ratings was .99, 95% CI [.981, .999], indicating excellent reliability among participants. Refer to Appendix B. for the SPSS (IBM, Armonk, NY) output.

The relationship between DUC scores and participants' perceptions of device complexity appeared non-linear (Figure 17). Therefore, the researcher ran a curve estimation regression analysis to determine the model that fit the data best. The results indicated that a logarithmic trend was an optimal fit, R^2 = .98. The proportion of variance in DUC scores accounted for 98.3% of the variance in participants' complexity ratings with adjusted R^2 = 97.9%, a high size effect according to (Cohen, 1988). The positive value indicates that the variables have a positive relationship. Therefore, participants rated devices with higher DUC scores as more complex.
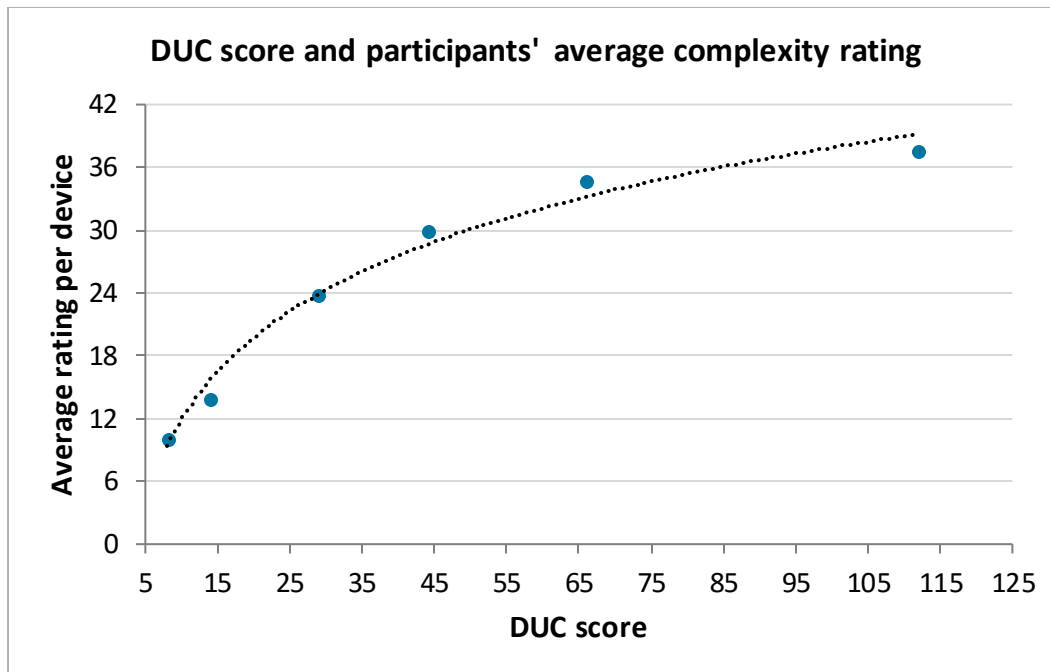
Figure 17: DUC score and participants' average complexity rating per device

The variables' logarithmic relationship suggests that once a device reaches a particular complexity level, small incremental differences in complexity become increasingly difficult to detect. These findings align with Weber's Law, which expresses the general relationship between the intensity level of a stimulus and the minimum amount the stimulus' intensity must be altered for a person to perceive a change. The phenomenon is commonly referred to as the difference threshold. According to Weber's Law, the difference threshold is a constant proportion of the stimulus' initial intensity level. For example, lighting a candle in a dark room causes a noticeable increase in illumination level. However, a difference in illumination level is less noticeable when a candle is lit in a room containing 150 burning candles. Small changes in intensity are difficult to detect when the baseline intensity level is relatively high. Thus, participants were more

41

sensitive to a certain increase in complexity when device complexity was

relatively low and less sensitive to the same increase in complexity when the

device's overall complexity was higher.

6. Primary Study

6.1. Research Questions

The primary study examined the relationship between child development and a child's readiness for participating in the administration of his or her health care. The researcher predicted that children who are ten years of age and older are capable of demonstrating task mastery with home health care devices.

The researcher gauged children's task mastery by comparing their performance to that of adults. The primary study explored the impact of child development variables on use error rate. The researcher hypothesized that as children develop biologically (i.e., hand size, grip strength, dexterity) and cognitively, (i.e., reading ability, working memory capacity, understanding of numerosity) use error rate decreases. In addition, the primary study related device complexity to the prevalence of use errors. The researcher hypothesized that there is a positive relationship between DUC scores and use error rate.

6.2. Method

6.2.1. Participants

The primary study included fifty-eight participants from two distinct groups: adults and children. The researcher recruited Tufts University undergraduates to represent the adult user group. Four adult females and three adult males participated in the study. The adult group consisted of seven individuals ranging in age from 19 to 21 years (M = 20.30 years, SD = .75 years).

Although the FDA defines an adult as 22 years of age or older, the researcher classified Tufts students as adults due to prior study results that showed individuals nearing the FDA's pediatric upper age limit received relatively high scores on cognitive and motor ability tests (Center for Devices and Radiological Health, FDA, n.d.; Hodes, et al., 2014; Tulsky et al., 2013; Wang et al., 2013).

The researcher recruited the adult participants from Human Factors courses. The researcher required them to have at least a Sophomore standing and 3.5 cumulative grade point average (GPA) to facilitate the assessment of child performance relative to individuals who demonstrated superior academic ability.

Fifty-one children participated in the study. The researcher recruited child participants from the Medford, Arlington, and Somerville communities. The child group consisted of 26 males and 25 females ranging in age from six to 12 years (M = 9.51 years, SD = 2.04 years). Moreover, the study included seven to eight child participants at each age.

The study screened for individuals receiving or eligible for special education services (e.g., Individualized Education Plan) or a 504 Plan. Additional exclusion criteria included individuals who did not speak English fluently or had experience using a blood glucose meter or nebulizer.

Each participant received a $30 Amazon gift card for participating in the study. Adult participants received the additional compensation of extra course credit for their participation.

6.2.2. Functional Metrics

The literature identified several biological and cognitive factors that have a potential impact on use error rate. This section describes how the researcher measured and scored each developmental variable.

6.2.2.1. Hand Length

Description: The researcher collected anthropometric data on participant hand length.

Procedure: Participants placed their dominant hand on a piece of paper. The researcher marked either side of the hand's base and traced the hand.

Scoring: The researcher measured the distance from the hand's base to the middle finger's tip in centimeters using a tape measure at a later point in time.

6.2.2.2. Grip Strength

Description: The researcher measured participant grip strength using a dynamometer (Figure 18).

Figure 18: Jamar® Plus Digital Dynamometer

Procedure: Participants wrapped their dominant hand around the dynamometer's handle, and positioned their arm at a right angle against the trunk. With the wrist in a neutral position, participants squeezed the handle with maximum force. Participants repeated the procedure with their non-dominant hand.

Scoring: The dynamometer provided a value indicating the amount of force participants were able to produce in pounds. The researcher recorded the strength of participants' dominant and non-dominant hands.

6.2.2.3. Dexterity

Description: The researcher administered the NIH Toolbox 9-Hole Pegboard Dexterity Test to assess participants' manual dexterity. Reuben et al. (2013) conducted a study that confirmed the test's validity. The pegboard is shown in Figure 19.
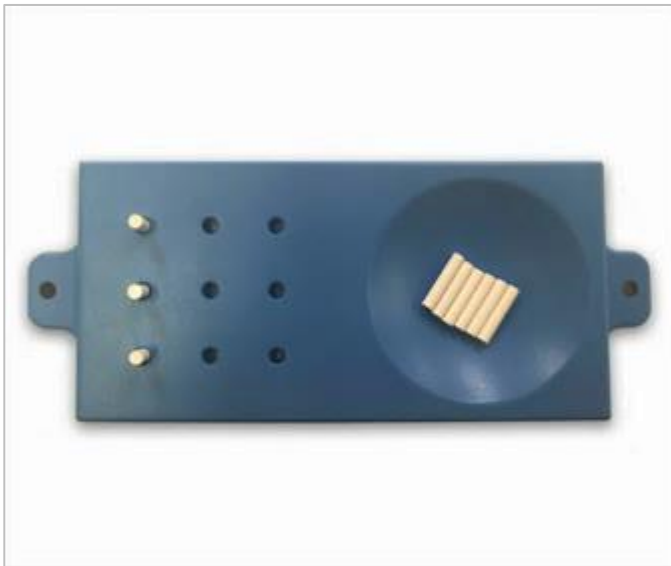
Figure 19: Jamar® 9-Hole Peg Test Kit

Procedure: Participants placed nine pegs into the pegboard's holes one at a time using only their dominant hand. When all nine holes were filled, participants removed the pegs one at a time from the pegboard. The researcher prompted participants to complete the task as quickly as possible. Participants completed one practice trial and one timed trial. Then, participants repeated the procedure with their non-dominant hand.

Scoring: The researcher recorded the amount of time it took participants to complete the task in seconds. The researcher only recorded data from the second trial with each hand.

6.2.2.4. Working Memory

Description: The researcher measured the working memory of participants using the NIH Toolbox List Sorting Working Memory Test. Tulsky et al. (2013) confirmed that the test is a valid measure of working memory capacity.

Procedure: The test requires participants to remember a series of items and then recite them in the correct order. The researcher used an iPad 2 to administer the test. The test's first task required participants to immediately recall and sequence a series of animals that were presented on the screen one at a time. For each object, the iPad displayed a picture, and provided the object's name orally (i.e., played an audio recording) and visually (i.e., written text). Participants were instructed to recite the objects is size order from smallest to largest. Each food list increased by one food item until the participant responded to two consecutive lists incorrectly.

The second task included both food and animals in each series. Participants were instructed to recite the food first in size order from smallest to largest, followed by the animals in size order smallest to largest. The test ended when participants responded to two consecutive lists incorrectly.

Scoring: Participants received points for recalling and ordering the lists correctly. The score ranges from 0-26.

6.2.2.5. Reading Ability

Description: To assess the participants' reading ability, the researcher used the San Diego Quick Assessment of Reading Ability (SDQA) (La Pray & Ross, 1969). SDQA measures participants' ability to recognize words out of context, also known as isolated word recognition. Smith Jr and Harrison (1983) conduct a study comparing SDQA to two widely used graded word reading tests and found that the scores were statistically significantly correlated with SDQA scores. Therefore, Smith Jr and Harrison (1983) deemed SDQA a reasonable alternative for rapidly determining reading level estimates.

SDQA consists of 13 graded word lists, with ten words in each list. La Pray and Ross (1969) designed the metric such that each list represents a distinct grade level spanning from pre-primer (i.e., preschool) to eleventh grade.

Procedure: The researcher presented participants with a word list two grades below the participant's grade level. Participants read the words from each list aloud. The researcher terminated the test when participants read three or more words incorrectly on a single list.

Scoring: The participant's reading level was the grade level associated with the last word list in which the participant read at least eight words correctly.

6.2.2.6. Functional Numerosity

Description: A functional numerosity test assessed participants' ability to read and sequence numerical values spanning from one to three digits.

Furthermore, the test evaluated the extent to which participants understood decimal notation.

Procedure: Participants read the five numbers aloud from an iPad screen. Then, the researcher instructed participants to sort the numbers from smallest to largest.

Scoring: Participants received points for reading and sequencing the numbers correctly. The score ranges from 0-5.

6.2.3. Survey

Participants responded to a paper-based survey. Parents of child participants completed the survey on their child's behalf. The researcher had three main objectives for distributing the survey. The first objective was to ensure participants met the study inclusion criteria. Participant responses enabled the researcher to confirm the accuracy of the information collected during the screening process, thereby ensuring the individuals qualified to participate in the study. The second objective was to elicit data requested by the NIH Toolbox iPad application. The application required the entry of certain information to administer the assessments, such as participant handedness. The researcher's third objective for distributing the survey was to collect responses to several questions that pertained to exploratory variables that the research team thought might affect use error rate.

6.2.4. Devices

All participants interacted with four home health care devices intended for use by the pediatric population. To increase the generalizability of the study results, the researcher assessed participant performance using devices with varying indications for use. Specifically, participants performed hands-on tasks with two different blood glucose meters (Figure 20 and Figure 21) and nebulizers (Figure 22 and Figure 23). Blood glucose meters are typically used by individuals diagnosed with diabetes to monitor the approximate level of glucose in their blood. Individuals diagnosed with a respiratory disease, such as asthma, chronic obstructive pulmonary disease (COPD), or cystic fibrosis commonly use nebulizers to administer their medication in the form of a breathable mist.

Figure 20: Nipro® SideKick Blood Glucose Testing System
(Blood glucose meter 1)

Figure 21: OneTouch® VerioIQ Blood Glucose Monitoring System (Blood glucose meter 2)

Figure 22: Philips® Respironics InnoSpire Elegance Compressor Nebulizer System (Nebulizer 1)

Figure 23: PARI® eRapid Nebulizer System (Nebulizer 2)

To assess the extent to which device complexity predicted use error rate, each device category (i.e., blood glucose meters, nebulizers) included one device with a relatively low and high level of complexity. The researcher utilized the DUC procedure to quantify device complexity (Appendix C.).

Importantly, the researcher modified the DUC procedure such that the functional trees were task-driven. The function trees excluded device components that did not pertain to the study tasks outlined in the following thesis section. For example, to ensure the devices did not lose power during study sessions, the researcher charged the devices before each session. As such, participants were not given the opportunity to interact with warning messages involving the device power levels, thereby eliminating the possibility of committing a use error related to that device component.

6.2.5. Tasks

The researcher employed FDA recommended medical device usability testing procedures to acquire use error rates (U.S. Department of Health and Human Services, FDA, 2016). Participants performed a series of hands-on tasks with four medical devices; two nebulizers and blood glucose meters. The researcher designed the primary study such that the medical device tasks simulated true use-case scenarios. The intent was to expose the wide range of use errors children might commit when placed under inopportune, yet realistic circumstances. Specifically, participants did not have prior experience using blood glucose meters or nebulizers or receive training before performing the medical device tasks. According to Wiklund, Kendler, & Strochlic (2016), designing the study to impose stress on participants increases the likelihood of uncovering dangerous use errors.

Notably, participants performed the tasks without parental support. Specifically, the researcher asked parents to limit their involvement during the session to simply observing the child. When the participant encountered an obstacle that prevented them from advancing to the next task step, the researcher provided assistance.

The medical devices included in the study were cleared/approved for market. However, participants were merely asked to manipulate the devices and simulate their use. The researcher did not instruct participants to inhale medicine through the nebulizers or lance themselves for glucose testing. The

devices' associated packaging and documentation (e.g., instructions-for-use [IFU]), quick reference guide [QRG]) were available to the participants for all tasks. To account for ordering effects on participants' performance with the devices, the researcher counterbalanced the order in which participants interacted with each device.

After administering the functional tests, the researcher oriented participants to each device, including brief descriptions of the devices' intended users and indications for use. Before working with each device, participants read a task description aloud from a card and then performed the task. The researcher read the task information to participants who did not demonstrate the literacy skills to read the material independently.

Participants performed the same task with both nebulizers. The task prompted participants to imagine they returned home from the pharmacy with the new device and the physician instructed them to take a 2.5 ml breathing treatment. The researcher provided participants with saline nebulizer solution (i.e., sodium chloride) to prepare the device for a treatment. Each participant utilized a new mouthpiece to simulate the inhalation to prevent cross-contamination by multiple users. The new mouthpiece was not attached to any other nebulizer pieces to ensure participants did not inhale any residual medication.

The task involving blood glucose meters instructed participants to take a blood glucose test, and then locate and read the last three blood glucose values

the meter saved. The researcher did not require participants to prick themselves with the lancing device or interact with any medicine. Rather, the participants simulated the blood glucose test with control solution.

To determine whether participants read and interpreted the glucose value correctly, the researcher prompted participants to (1) read the glucose value they attained, and indicate whether the glucose value was (2) good or bad, (3) too high or low, and (4) the action s/he would take next (e.g., take insulin, eat a snack). Notably, before performing blood glucose meter tasks, the researcher explained the diabetes-related concepts necessary to interpret the glucose value correctly. Moreover, participants had access to a chart that specified normal blood sugar levels for children with diabetes.

6.2.6. Interviews

After performing each task, participants responded to a series of interview questions. First, participants' rated the device's ease of use on a 7-point scale (Figure 24). The researcher color-coded the scale to increase the reliability of responses from children who had difficulty reading the scale's text.
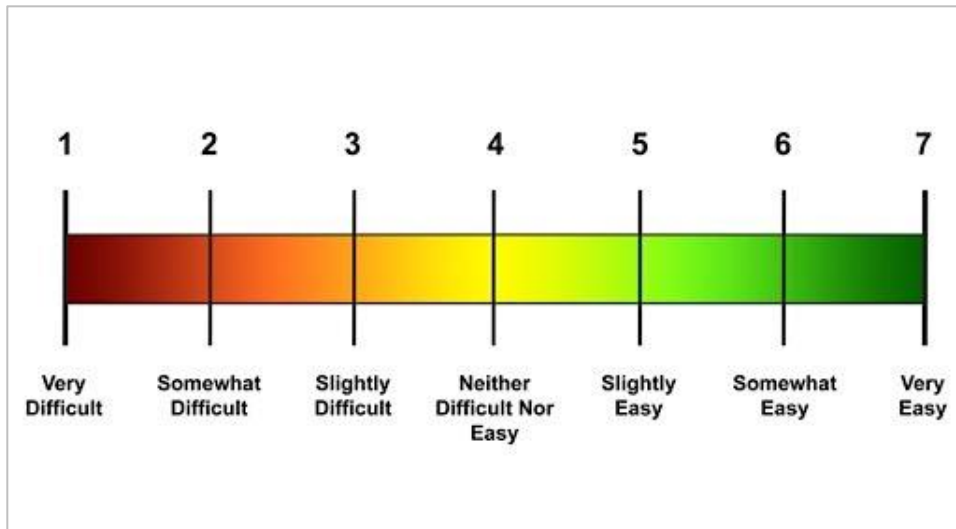
Figure 24: Ease of use scale

Then, the researcher asked participants to recall whether they made any mistakes (i.e., use errors) during the task. According to Wiklund et al. (2016), participants might commit use errors that are difficult for the researcher to detect, such as misinterpreting an on-screen menu option's label. When participants reported committing a use error during the primary study, the researcher asked follow-up questions to facilitate accurate documentation. In addition, the researcher asked participants to recall any times when they almost made a mistake, but ultimately did not (i.e., close call) or found a certain aspect of the task confusing or difficult (i.e., operational difficulty) (Wiklund et al., 2016).

After participants had performed tasks with both devices in a particular device category, the researcher followed up on any use errors, close calls, operational difficulties, or instances when the researcher needed to intervene and provide assistance. The interview's purpose was to ascertain the root causes

associated with the issues. The root cause analysis aided in differentiating

between interaction problems associated with personal developmental deficits

and those associated with underlying design issues. For example, a child might

commit a use error because the device's display suffers from low contrast.

Therefore, we can attribute the error to a usability issue that is not child-specific.

6.2.7. Experimental Protocol

During the recruiting process, the researcher provided participants with information regarding the study procedure and basic elements of informed consent (e.g., foreseeable risks, benefits, confidentiality of records). The researcher scheduled individuals (or parents of individuals) who expressed interest in (their child) participating in the study and met the study inclusion criteria. All study activities took place at Tufts University's Human Factors Usability Lab (Figure 25). Each study session lasted up to 1.5 hours.

Figure 25: Human Factors Usability Lab

When participants arrived at the usability lab, the researcher provided each parent and adult participant with a consent form and survey. Child participants received assent forms. Once the researcher addressed all questions concerning the consent and assent forms, the study activities began. The

researcher asked parents of child participants to remain in the usability lab for

the study session's duration. Notably, the researcher recorded each study session

to facilitate data entry. Study session activities included:

- Explaining the study's purpose.

- Administering functional tests.

- Orienting participants to each medical device, including brief descriptions of the devices' intended users and indications for use.

- Instructing participants to perform one hands-on task with each medical device.

- Collecting participants' subjective rating of each device's ease of use.

- Asking participants whether they recalled making any mistakes, incidences of close calls, or operational difficulties.

- Following up on any use errors, close calls, operational difficulties, or assists that were not addressed during the post-task interview.

- Compensating the participant.

6.2.8. Data Collection

As participants performed the study tasks, the researcher utilized

checklists to determine whether participants were taking the appropriate actions

with each device. The researcher generated the checklists using the perception,

cognition, and action (PCA) task analysis technique (International

Electrotechnical Committee (IEC), 2015). The procedure facilitated a more

comprehensive understanding of the task workflow and aided in identifying the

types of errors that could result in harm.

During approximately half of the study sessions, a research assistant recorded the information in a Microsoft® Excel spreadsheet simultaneously. The research team collected the following data:

- Use errors
- Close calls
- Operational difficulties
- Instances of assistant from the researcher
- Ease of use ratings
- Level of IFU usage

To ensure the research team was consistent in their assessments, the researcher developed guidelines to which the team adhered during the study sessions. Table 1 describes the criteria the researcher implemented to document the extent of participant IFU usage.

Table 1: IFU usage scoring criteria

IFU score	Description
1	The participant did not acknowledge the IFU (e.g., did not remove the IFU from its packaging).
2	The participant acknowledged the IFU (e.g., removed the IFU from its packaging, unfolded or opened the IFU), but did not reference the IFU's text or graphics.
3	The participant referenced the IFU's text or graphics, but did not read all IFU text that pertained to accomplishing the task.
4	The participant read all IFU text that pertained to accomplishing the task.

The researcher developed similar criteria for use errors, close calls, and operational difficulties. For example, the blood glucose meter IFU instructs users to wash the injection site before performing a blood glucose test to reduce the chance of infection. Therefore, the researcher definitively determined that participants who did not use the hand sanitizer provided or indicate that they would wash the injection site committed the use error. Furthermore, the research team documented participant reported root causes and any comments about their experience working with the devices.

Notably, the researcher collected data on device ease of use and IFU usage to ensure the primary study fully emulated a typical validation usability test design. The analyses involving these data are outside the scope of this thesis, and thus not addressed in the Results section.

6.3. Results and Discussion

6.3.1. Age and Use Error Rate

The first research objective pertained to evaluating the relationship between use error rate and age. The researcher sought to investigate the transition point when children become capable of participating in their own care. Table 2 provides descriptive statistics of use error rate categorized by age group and subgroup. In this section, use error rate refers to the average number of safety-related use errors children committed with all devices. The intent was to generate results that are generalizable to a wide range of home health care

devices on the market that are intended for use by children. Overall, use error rate appeared to decline with age. The scatter plot in Figure 26 represents the continuous relationship between children's age and use error rate.

Table 2: Use error descriptive statistics for groups and subgroups

Age groups	N	Range	Mean	SD
Children (6 – 12 years)	51	1.8 – 9.0	5.5	0.3
6 years	7	7.0 – 9.0	8.3	0.7
7 years	8	5.5 – 8.3	6.9	0.8
8 years	7	4.8 – 7.8	6.3	1.3
9 years	7	3.5 – 6.5	5.0	1.1
10 years	7	2.8 – 5.8	4.6	1.0
11 years	8	1.8 – 5.3	3.7	1.3
12 years	7	1.8 – 4.5	3.6	1.2
Adults (19 – 21 years)	7	2.3 – 4.0	3.3	0.7
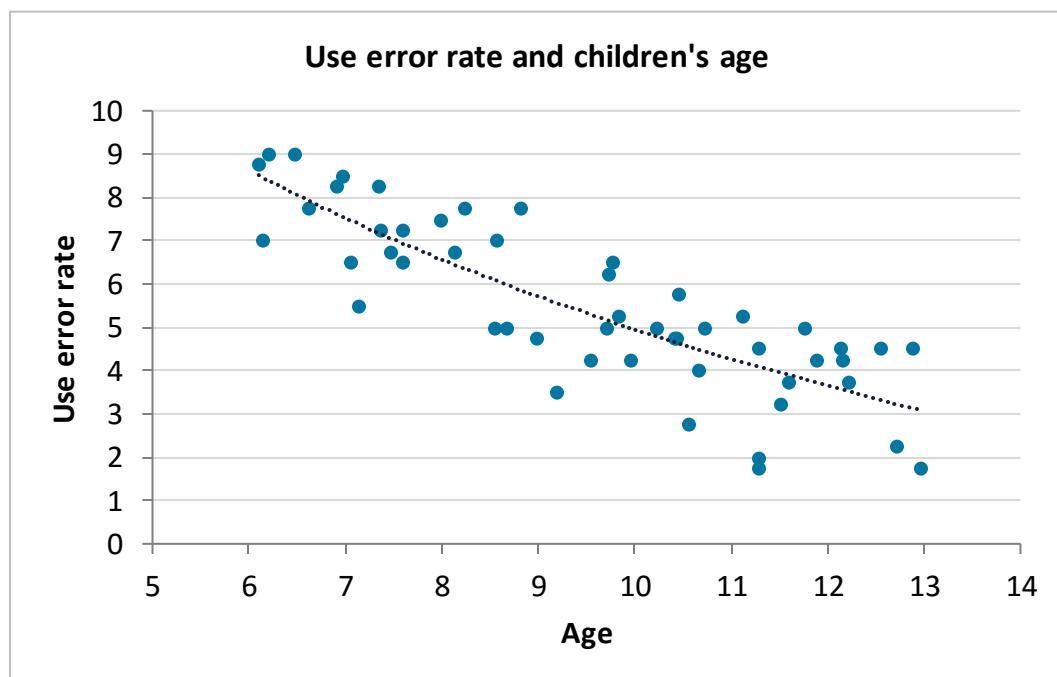
Figure 26: Use error rate and children's age

An initial visual inspection of the data's trend revealed a potential curvilinear relationship between the variables. Use error rate appeared to decrease steadily and reach a plateau at later ages. This triggered further analysis to determine the best-fit line. A curve estimation regression analysis confirmed that a logarithmic model optimized the fit. Therefore, the researcher transformed the x-axis (age) logarithmically to base 10 to coax the independent and dependent variables into a linear relationship. To determine the extent to which age and use error rate were related, the researcher computed a Pearson product-moment correlation coefficient. Refer to Appendix D. for the primary study SPSS output.

The correlation results indicate that there was a statistically significant negative relationship between the two variables, $r = -.84$, $p < .001$. Overall, age was strongly associated with the average number of use errors children committed with all devices. The variables' negative relationship indicates that children who were older were less likely to commit use errors. Furthermore, the variables' curvilinear relationship denotes little variation among older children relative to younger children regarding the number of use errors they committed. The result suggests that at a particular point in development, children converge in their ability. Therefore, the six to 12 age range might contain the minimum age required to perform medical device tasks with minimal error.

To identify the point in development when children possess the prerequisites to perform medical device tasks successfully, the researcher

compared children's performance to that of adults. The adult group included

Tufts University students who possessed high cognitive and physical ability. The

intent was to identify the minimum number of use errors the adult population is

likely to commit with each medical device.

Table 3 describes the use errors that adults and child participants

committed during the study. The table excludes use errors that were unique to a

particular device.

Table 3: Number and percentage of participants who committed particular use errors (UEs)

Use error description	No. of adults who committed UE (%)	No. of children who committed UE (%)
Blood glucose meters		
Did not check the expiration date	7 (100)	51 (100)
Did not wash hands	6 (85.7)	50 (98.0)
Did not close the vial cap immediately	3 (42.9)	30 (58.8)
Did not touch sample to the test strip's channel	2 (28.6)	33 (64.7)
Did not discard the test strip properly	1 (14.3)	13 (25.5)
Did not interpret the test result correctly	0 (0)	20 (39.2)
Sample was too small	0 (0)	17 (33.3)
Did not select a correct injection site	0 (0)	16 (31.4)
Applied the sample to the test strip before inserting the test strip into the meter's port	0 (0)	10 (19.6)
Read the blood test result incorrectly	0 (0)	9 (17.6)
Inserted the test strip in the wrong orientation into the meter	0 (0)	5 (9.8)
Removed more than one test strip from the vial	0 (0)	4 (7.8)
Nebulizers		
Did not inspect for damage	6 (85.7)	51 (100)
Dispensed too much solution	4 (57.1)	30 (58.8)
Did not secure lips around the mouthpiece	2 (28.6)	13 (25.5)
Did not dispense solution into the medicine cup	1 (14.3)	28 (54.9)
Dispensed too little solution	1 (14.3)	12 (23.5)
Spilled solution	1 (14.3)	10 (19.6)
Did not breathe through the mouthpiece	0 (0)	22 (43.1)
Did not sit upright during the treatment	0 (0)	9 (17.6)

Note. Participants might have committed a particular use error one or more times with one or both devices.

The scatter plot in Figure 27 compares the average number of use errors children and adults committed with all devices. The area between the dotted

lines represents the 99% confidence interval for mean adult use error rate, 99% CI [2.336, 4.306]. The confidence interval indicates the range estimated to contain the population mean for adult participants. The researcher selected a slightly higher confidence level than the typical 95% level to increase the likelihood that the interval contained the true mean. The researcher defined the 99% confidence interval upper limit as an estimation of the maximum number of use errors adults are likely to commit when performing tasks with medical devices similar to those in the study.
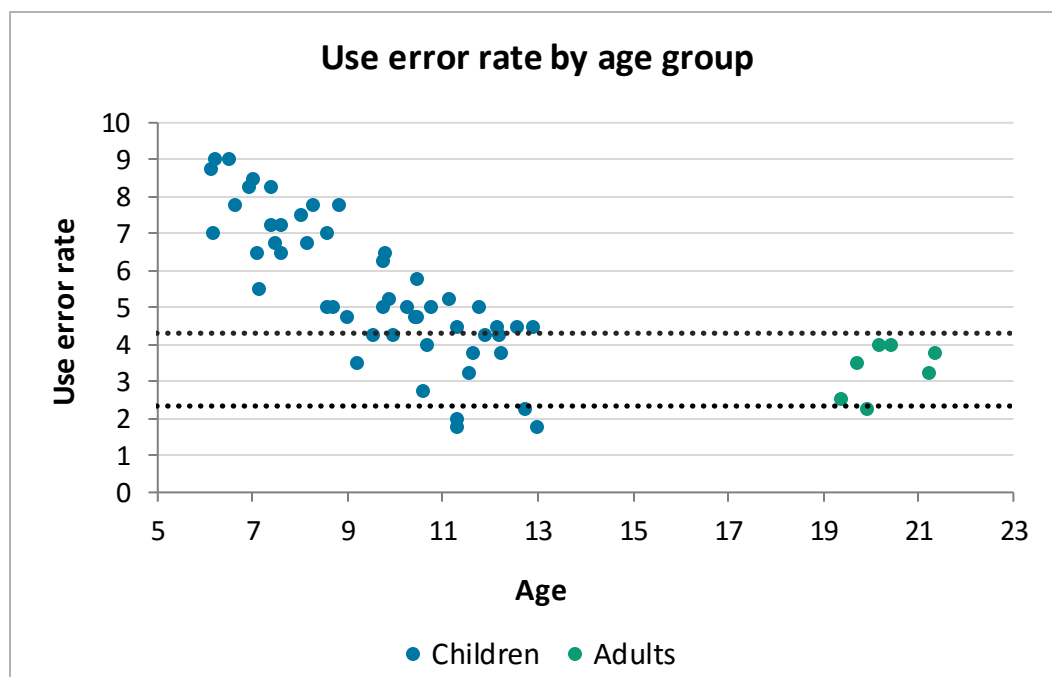
Figure 27: Use error rate by age group. The area between the dotted line indicates the 99% CI for adult use error rate.

Figure 28 represents the percentage of children at each age who performed the medical device tasks with minimal error. The researcher defined

67

minimal error as falling below the 99% confidence interval upper limit for adult

participants.
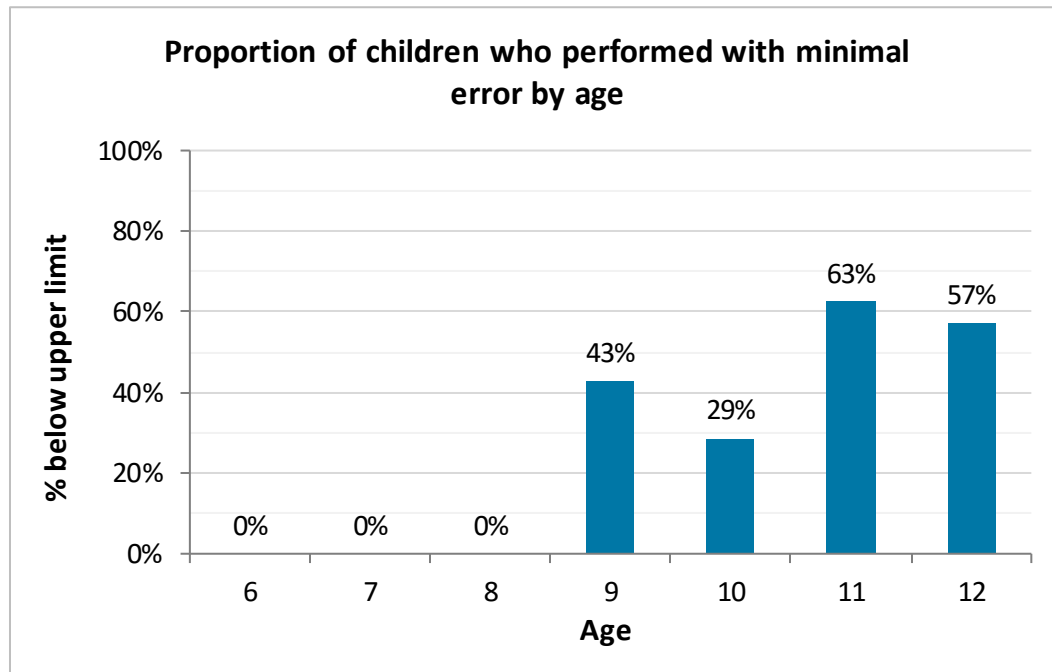
Figure 28: Percentage of children who performed with minimal error by age

All participants who were eight years of age or younger fell below the

interval upper limit. Those children committed a high number of use errors

compared to adult participants. Forty-three percent of children who were nine

years of age performed with minimal error. The substantial rise suggests there

were 9-year-olds who had met the developmental milestones necessary to use a

medical device successfully. It appeared there were additional variables creating

a barrier to sufficient performance for children who were younger than nine.

The percent decreased slightly in children who were ten years of age;

twenty-nine percent performed with minimal error. The drop in performance is

likely due to the individual differences between children who were the same age,

noting that the study design was cross-sectional, rather than longitudinal. The random sample of nine-year-olds might have been more developmentally equipped to perform the medical device tasks than those who were ten years of age. Moreover, the small sample sizes likely inflated the minor discrepancies between children close in age.

More than half of children who were 11 and 12 years of age performed with minimal error. Although this is high compared to the sample of younger children, roughly 40% of 11- and 12-year-olds committed an excessive number of use errors. As such, children's developmental capabilities might play a significant role in their ability to use medical devices adequately. These results align with the previous notion that children might not converge in their ability at the same age. Therefore, the researcher sought to investigate the developmental milestones the children who performed well had met.

The following subsections assess the relationship between various external factors and use error rate. Understanding the degree to which each variable was related to use error rate facilitated a more accurate interpretation of the factors that had the strongest association with task performance. The researcher categorized the data collected from the functional assessments and survey into two variable groups: (1) developmental and (2) exploratory.

6.3.2. Developmental Variables and Use Error Rate

The literature indicated several developmental variables that might have a strong relationship with use error rate. The researcher measured two higher-

level aspects of child functioning, including biological and cognitive. This section

discusses the extent to which use error rate was related to the following

developmental variables:

- Biological
  - Hand length
  - Dominant and Non-dominant dexterity
  - Dominant and Non-dominant grip strength

- Cognitive
  - Working memory
  - Reading ability
  - Functional numerosity

The study protocol included valid metrics to assess children's level of

biological and cognitive development, each of which is described in the Method

section. Table 4 provides descriptive statistics of children's functional assessment

results.

Table 4: Descriptive statistics of children's functional assessment results

Developmental variable	Range	Mean	SD
Biological			
Hand length (cm)	12.3 – 18.8	15.2	1.5
Dominant dexterity (seconds)	16.7 – 26.2	21.0	2.6
Non-dominant dexterity (seconds)	16.7 – 35.9	23.6	4.2
Dominant grip strength (lbs)	14.4 – 67.0	33.5	12.3
Non-dominant grip strength (lbs)	14.6 – 67.3	31.6	11.3
Cognitive			
Working memory (number correct)	8 – 23	16.2	3.4
Reading ability (grade level)	-2 – 11	5.3	3.4

Functional numerosity (number correct)	0 – 5	3.3	1.8

The scatter plot in Figure 29 illustrates the relationship between hand length and the number of use errors children committed with all devices. The green data points represent children who perform the tasks with minimal error.
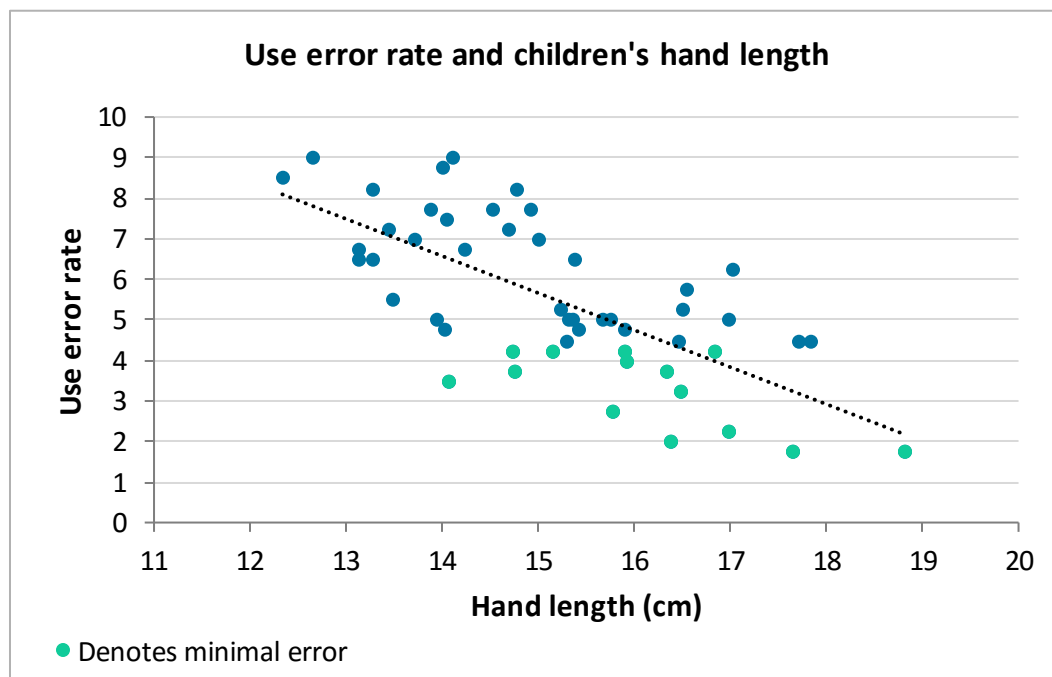
Figure 29: Use error rate and children's hand length

Thirty-six percent of children who had a hand length of at least 14.06 cm performed with minimal error. Therefore, possessing a hand length below approximately 14 cm (5.5 inches) might create a barrier to performing autonomously with medical devices. To confirm the accuracy of this result, the researcher computed a Pearson correlation to assess the strength of association between children's hand length and use error rate.

Hand length was significantly correlated with the average number of use errors children committed with all devices, $r = -.696$, $p < .001$. The negative value indicates that children with relatively longer hands were less likely to commit use errors. However, most developmental variables tend to correlate with age. Thus, children's age might have influenced the significant result. This triggered further analysis to determine the unique variance between hand length and age. The researcher ran a partial correlation and controlled for age to eliminate its impact on the result. Hand length and use error rate were not significantly correlated while controlling for age, $r = .023$, $p = .436$. Therefore, the researcher cannot conclude that children's hand length influenced use error rate.

The researcher ran an additional Pearson correlation to determine whether the other developmental variables were highly associated with use error rate and age. Notably, the researcher conducted a Spearman correlation to evaluate the extent to which functional numerosity scores were related to age because the data were not normally distributed.

The results indicated that all developmental variables were significantly correlated with use error rate as well as age (Table 5). Therefore, the preceding correlation analyses assessing the relationship between developmental variables and use error rate, control for the effects of age. Running partial correlations enabled the researcher to examine the unique influences of developmental variables on use error rate.

Table 5: Correlation between developmental variables and age

Developmental variables	Use error rate Correlation coefficient	Age Correlation coefficient
Hand length	-.696**	.848**
Dominant grip strength	-.737**	.871**
Non-dominant grip strength	-.712**	.832**
Dominant dexterity	-.411*	-.541**
Non-dominant dexterity	-.596**	-.682**
Working memory	-.685**	.690**
Reading ability	-.751**	.785**
Functional numerosity	-.795**	.884**

Note. * $p < .01$; ** $p < .001$

As discussed previously, children's grip strength might influence their ability to perform some home health care device tasks. Medical devices tend to require a relatively high amount of grip strength. For example, during the primary study, children likely employed grip strength when opening vial caps for glucose testing and connecting nebulizer components. As such, the researcher utilized a dynamometer to assess the amount of force participants were able to produce with their hands. The scatter plots in Figure 30 and Figure 31 represent the relationship between use error rate and children's dominant and non-dominant grip strength, respectively.
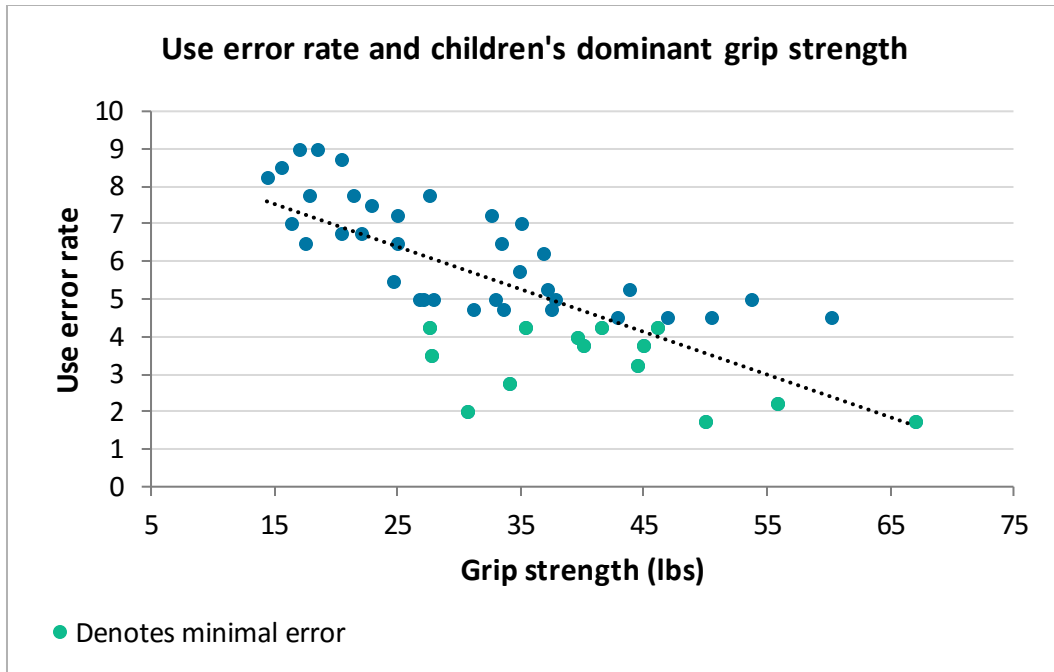
Figure 30: Use error rate and children's dominant grip strength
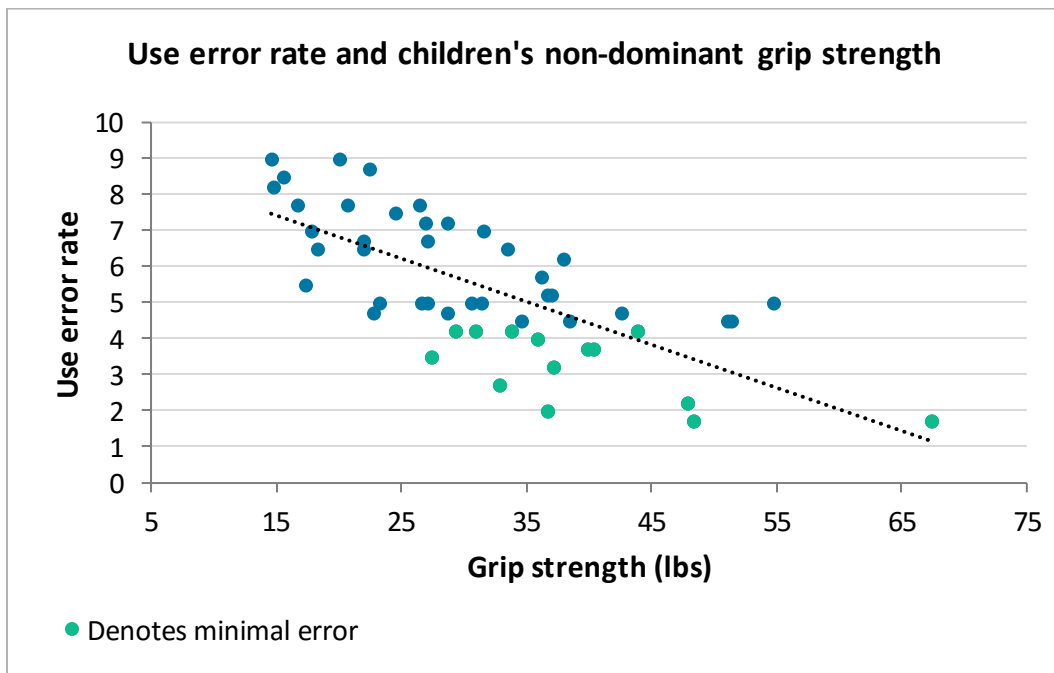
Figure 31: Use error rate and children's non-dominant grip strength

When the variance from age was eliminated, use error rate was not

statistically significantly associated with children's dominant, $r = -.054$, $p = .356$,

or non-dominant grip strength, *r* = -.070, *p* = .316. Similar to children's hand size, age appeared to control the relationship between grip strength and performance. Therefore, the researcher cannot conclude that children's grip strength was related to committing use errors. Subsequently, the researcher did not include an analysis concerning the level of grip strength children who performed with minimal error possessed.

The researcher hypothesized that children's level of fine motor ability was significantly related to use error rate. Both school nurses expressed that children with poor dexterity are less likely to perform medical device tasks successfully. Moreover, the literature indicated that dexterity tends to plateau during childhood, suggesting that children's age might not have a strong influence on the relationship between fine motor skills and committing use errors.

The researcher administered the 9-Hole Pegboard test to assess children's level of fine motor skill. The scatter plots in Figure 32 and Figure 33 illustrate the association of use error rate and children's dominant and non-dominant dexterity, respectively. The values on the x-axes represent the total number of seconds it took children to complete the task. Higher values indicate lower levels of fine motor ability.
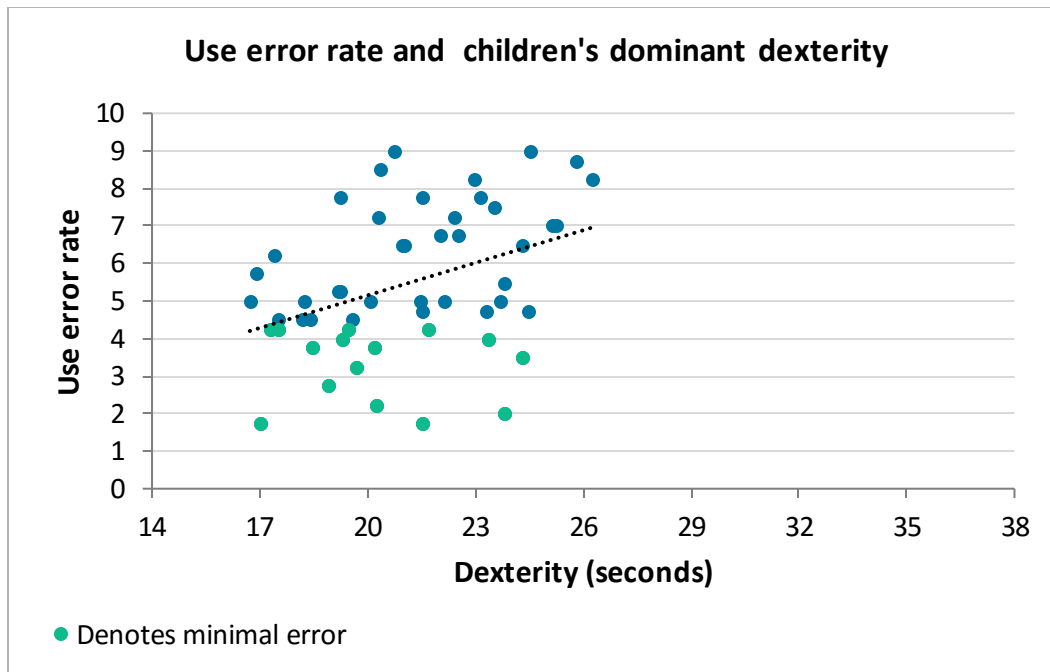
Figure 32: Use error rate and children's dominant dexterity
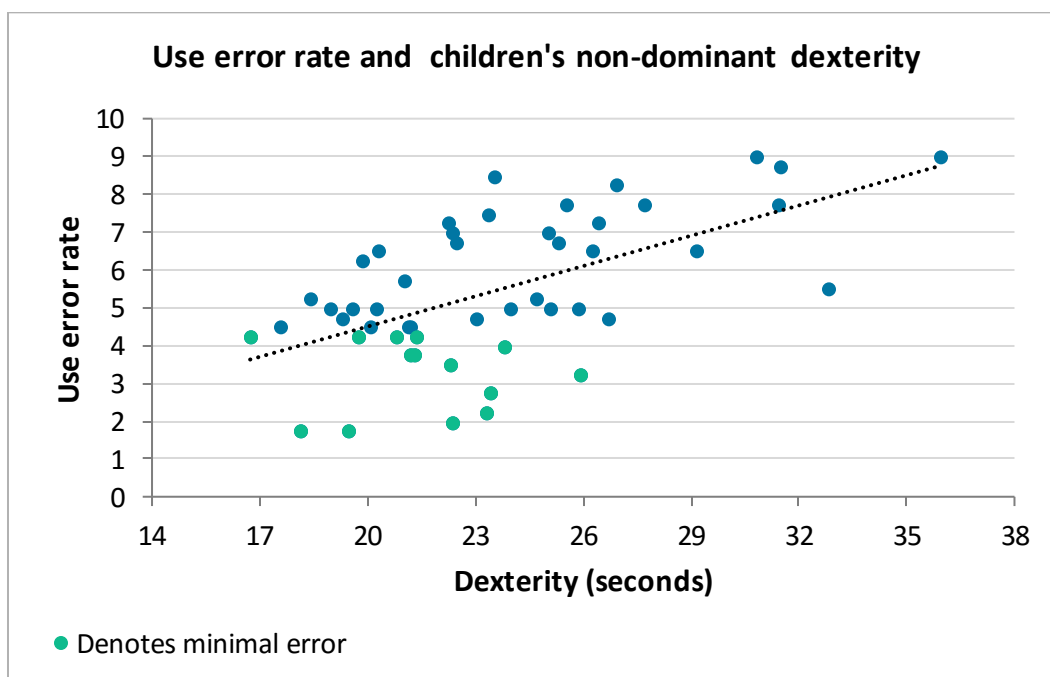
Figure 33: Use error rate and children's non-dominant dexterity

The partial correlation results indicated that use error rate was not significantly correlated with dominant, r = -.079, p = .292, or non-dominant dexterity, r = -.075, p = .303, while controlling for age. Although the scatter plots

might depict a relationship between the variables, the analysis revealed that age influenced their association greatly. The researcher cannot conclude that children's level of dexterity had an impact on performing the tasks. Overall, all biological variables were not significantly related to committing a greater number of use errors.

The researcher hypothesized that use error rate would decrease as children's working memory capacity increased. The literature indicated that developmental increases in working memory tend to have a positive effect on children's ability to communicate and problem solve effectively. The fact that children's working memory capacity typically plateaus around ten years of age lead to the prediction that working memory capacity impacts children's ability to perform medical device tasks properly.

Figure 34 represents the relationship between children's working memory score and use error rate. The values on the y-axis represent the score participants received on the working memory test. A higher score signifies a greater working memory capacity. The researcher ran a partial correlation to assess the variables' relationship.
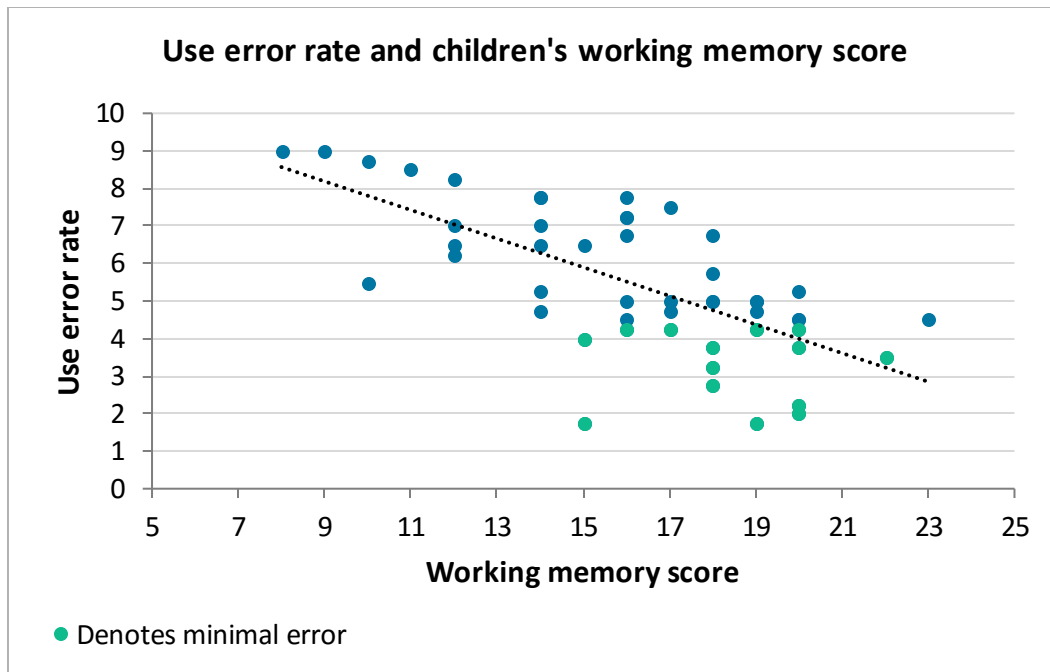
Figure 34: Use error rate and children's working memory score

There was a negative partial correlation between working memory score and use error rate while controlling for age, which was statistically significant, $r = -.279$, $p = .026$. The negative relationship indicates that children who were capable of recalling and sequencing more pieces of information were less likely to commit use errors. Importantly, age had little influence in controlling the relationship between working memory and use error rate. This prompted further analysis to ascertain the minimum working memory score children who performed with minimal error obtained.

The results indicate that approximately 40% of children who scored a 15 or higher on the working memory test committed the minimum number of errors while performing the medical device tasks. However, the working memory scores are rather arbitrary. Therefore, the researcher rescored the assessment to

produce scores that represent the number of information elements (e.g.,

animals, foods) children were able to recall and sequence correctly. The score

takes the average of the number of items children recited correctly on both tasks.

The scatter plot in Figure 35 shows the relationship between the two scoring

methods.

Raw vs. modified working memory score
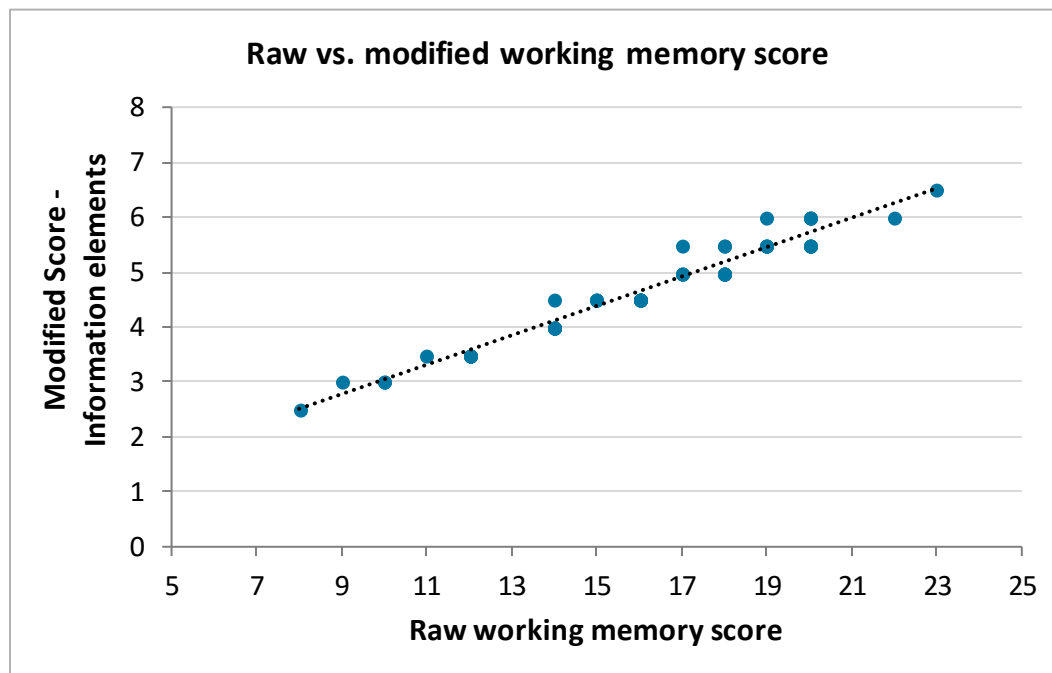
Figure 35: Raw vs. modified working memory score

The results of a Pearson correlation denoted a strong correlation between

the raw and modified working memory scores, $r = .976$, $p < .001$. Ultimately, the

modified score afforded a more practical interpretation of receiving a raw score

of 15 on the working memory test. A score of 15 or higher equates to reciting

four to five pieces of information correctly with the one- and two-category list,

respectively.

Furthermore, the researcher hypothesized that there would be a negative relationship between children's reading ability and use error rate. Increases in reading ability are known to influence children's capacity to process and understand information. Therefore, reading at a higher level might increase the likelihood of performing medical device tasks successfully. The researcher sought to evaluate the relationship between children's reading ability and use error rate. The scatter plot in Figure 36 illustrates the relationship between the variables. The reading ability assessment produced scores that indicated the grade level for which children read. The values on the x-axis denote the extent of children's reading level by grade level.
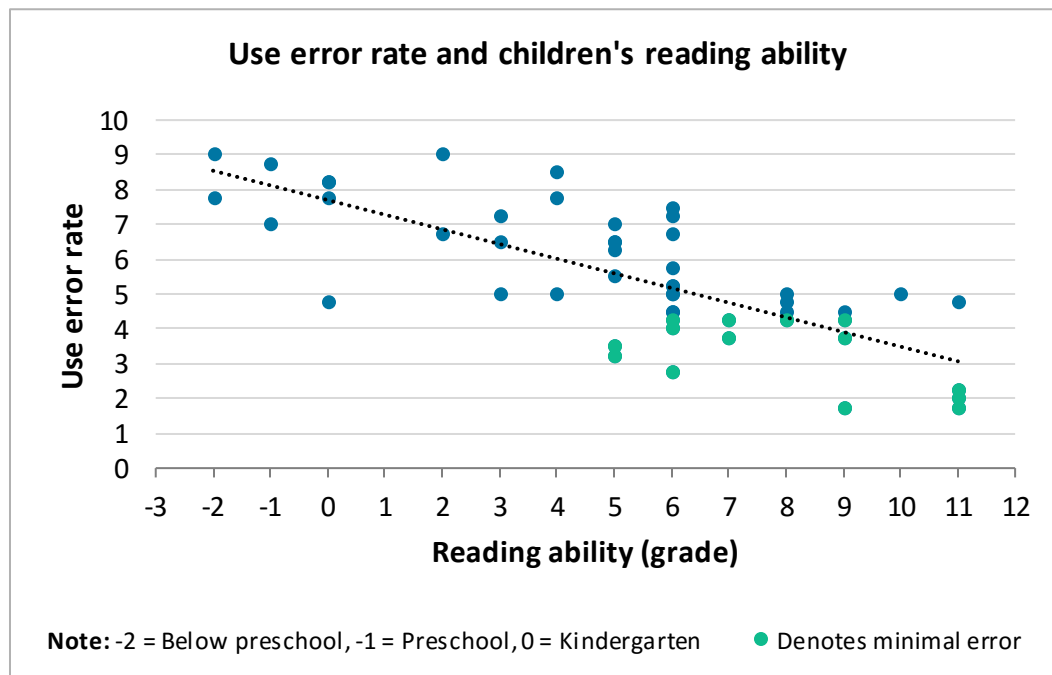
Figure 36: Use error rate and children's reading ability

The researcher ran a partial correlation to measure the magnitude of association between children's reading ability and use error rate, while

controlling for age. The partial correlation yielded a statistically significant

results, r = -.291, p = .020. Therefore, children who possessed a greater capacity

to read words out of context committed few use errors. Importantly, 40% of

children who performed the medical device tasks with minimal error read at a

fifth-grade level or higher. A substantially higher percentage of children, roughly

60%, performed at an effective level when they read at a seventh-grade level or

higher.

The final developmental variable the researcher tested for was children's

understanding of numerosity. The assessment evaluated participants' ability to

read and sequence numerical values spanning from one to three digits.

Furthermore, the test evaluated the extent to which participants understood

decimal notation. Many home health care devices require such skills to perform

tasks effectively. In the primary study, the blood glucose meters required the

ability to perform numeric operations with three digit numbers to interpret the

results correctly. Moreover, the nebulizers required measurement skills for

calculating medication doses. Thus, the researcher hypothesized that children

who received relatively high scores on the functional numerosity test would

commit fewer use errors (Figure 37).
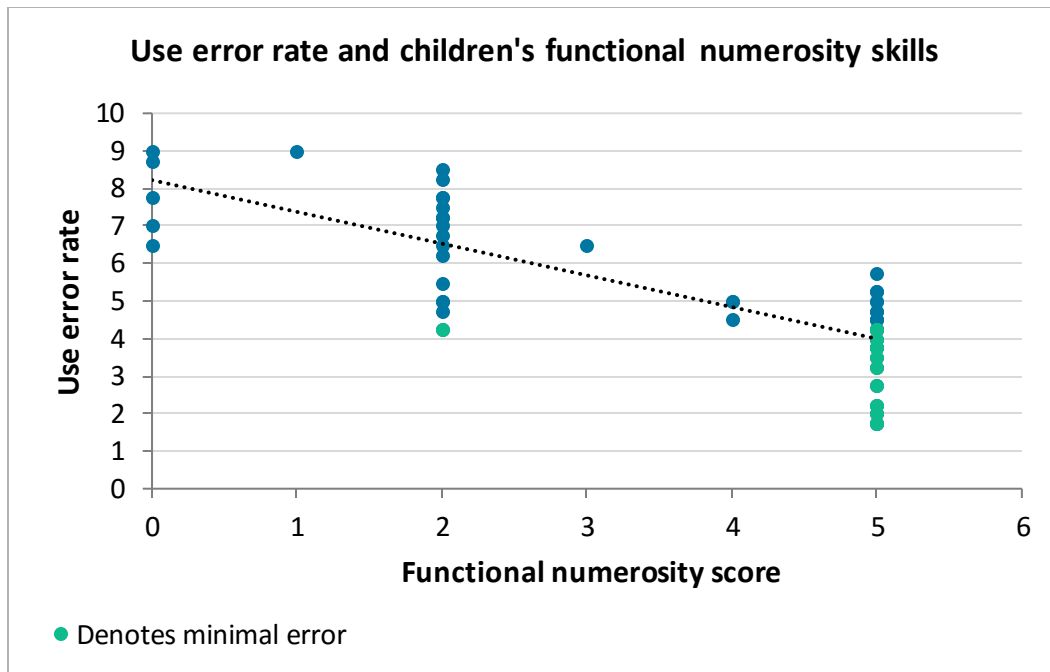
Figure 37: Use error rate and children's functional numerosity score

The researcher ran a partial correlation to determine the extent to which children's understanding of numerosity is related to use error rate when removing the effects of age. The results showed a non-significant correlation, $r = -.231$, $p = .057$. Functional numerosity scores fell just short of significance. Notably, the correlation evaluated the extent to which the variable was associated with committing use errors with all four medical devices. While numerosity was not statistically significantly related to the overall rate of use errors, the near-significant p-value suggests a correlation between children's understanding of numerosity and committing individual types of use errors. Therefore, the researcher ran a point-biserial partial correlation to assess the relationship between children's functional numerosity scores and committing discrete use error types, while controlling for age.
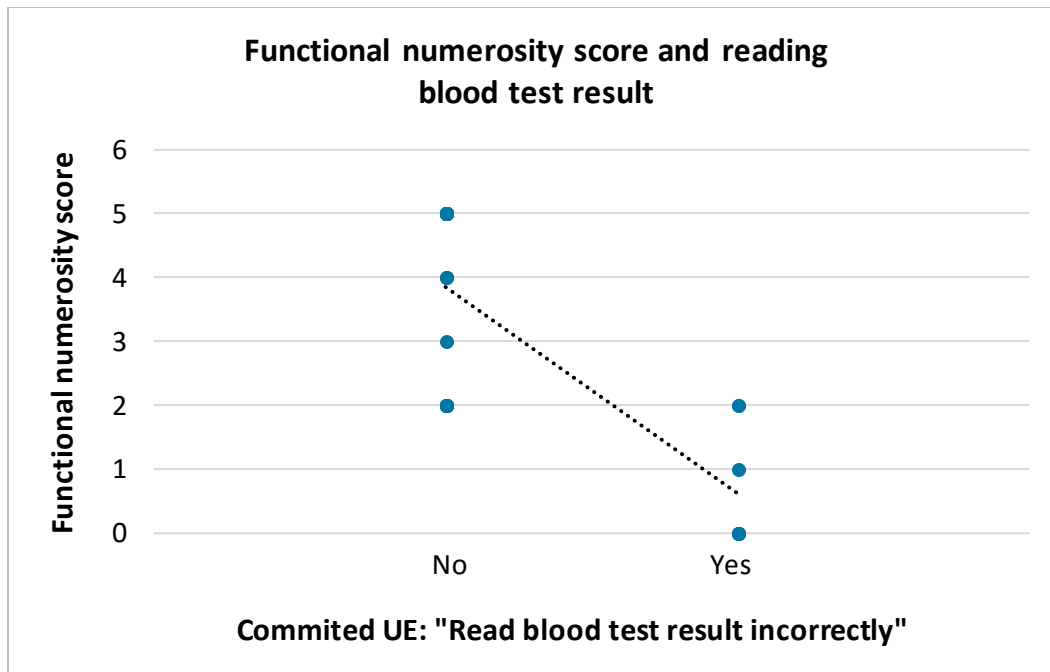
Figure 38: Children's functional numerosity score and reading blood test result

The analysis yielded a statistically significant result for the use error pertaining to reading the blood test result, r = -.293, p = .044. Children who were incapable of demonstrating a comprehensive understanding of numerosity tended to read the blood test result incorrectly. Although numerosity might not have a strong influence on children's overall task performance, the results indicated that relatively low numerosity scores were related to committing a critical error, and ultimately created a barrier to essential performance for those children.

6.3.3. Exploratory Variables and Use Error Rate

The survey elicited data on several exploratory variables that might

impact use error rate, including:

- Gender
- Handedness
- Parent education level
- Medical device familiarity

The researcher conducted a correlation analysis to assess the extent to

which each exploratory variable was related to use error rate. Table 6

summarizes the correlation results.

Table 6: Correlation between exploratory variables and use error rate

Cognitive variable	Correlation coefficient
Gender	-.021
Handedness	-.076
Parent education level	-.125
Medical device familiarity	-.043

There were 26 male and 25 female child participants in the primary study.

Male use error rate (M = 5.53, SD = 1.79) was slightly higher than female use

error rate (M = 5.45, SD = 2.09). Figure 39 summarizes the relationship between

gender and the average number of use errors children committed with all four

devices.
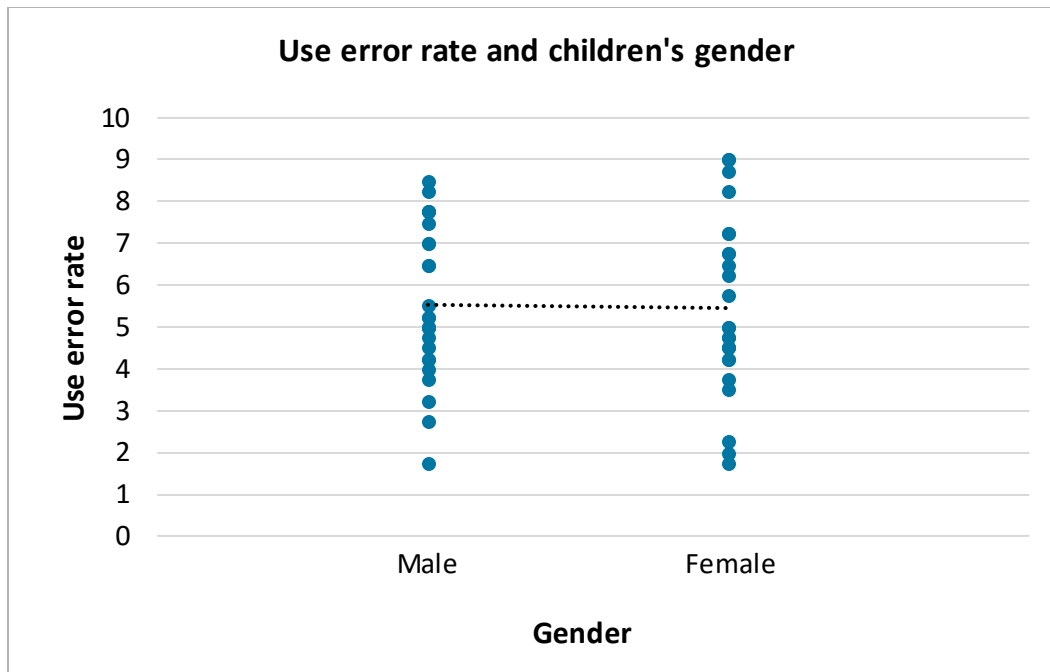
Figure 39: Use error rate and children's gender

To determine the strength of the relationship between children's gender and use error rate, the researcher ran a point-biserial correlation. The results indicated that there was not a statistically significant correlation between gender and use error rate, $r = -.021$, $p = .885$. Therefore, the researcher cannot conclude that gender influenced the number of use errors children committed.

Forty-six child participants were right-handed and five were left-handed. Left-handed participants ($M = 5.05$, $SD = .99$) committed slightly fewer use errors than right-handed participants ($M = 5.54$, $SD = 2.00$) on average with all four medical devices (Figure 40).

Use error rate and children's handedness
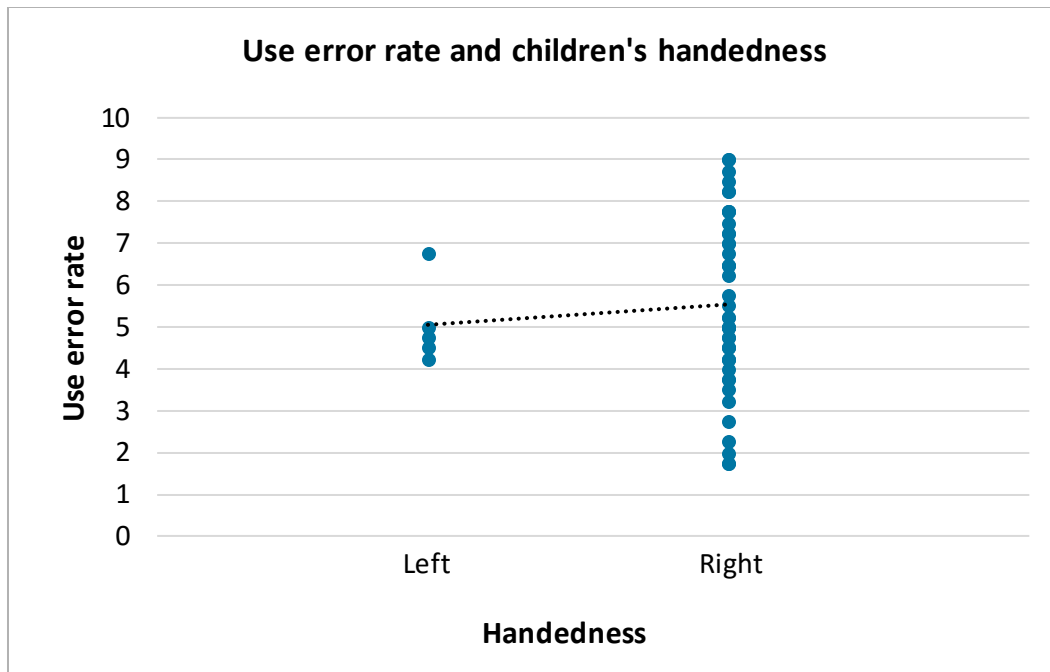
Figure 40: Use error rate and children's handedness

The results of a point-biserial correlation indicated that the relationship between use error rate and children's handedness was non-significant, $r = -.021$, $p = .595$. Differences in handedness did not appear to correlate with the number of use errors children committed.

The survey asked parents to indicate the highest level of education they completed. A majority of parents had received a bachelor or master's degree. Several parents reportedly had a doctorate, and three parents had solely a high school degree.
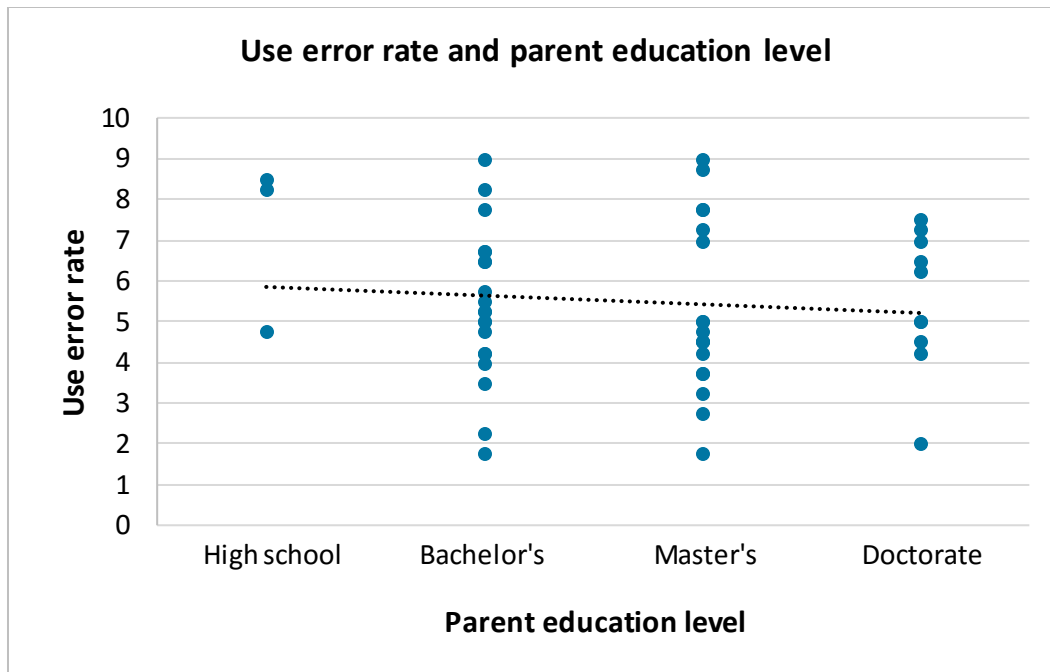
Figure 41: Use error rate and parent education level

The researcher ran a Pearson correlation to evaluate the relationship between the highest level of education participants' parents completed and the average number of errors children committed. The results indicated that the variables were not significantly related, $r = -.125$, $p = .383$.

Several children had seen their family members use medical devices in the past. Parents of participants reported a wide range of devices types, including nebulizers, blood glucose meters, inhalers, auto-injectors, and blood pressure gauges. The researcher conducted a Pearson correlation to determine whether children's familiarity with medical devices was related to committing use error (Figure 42).
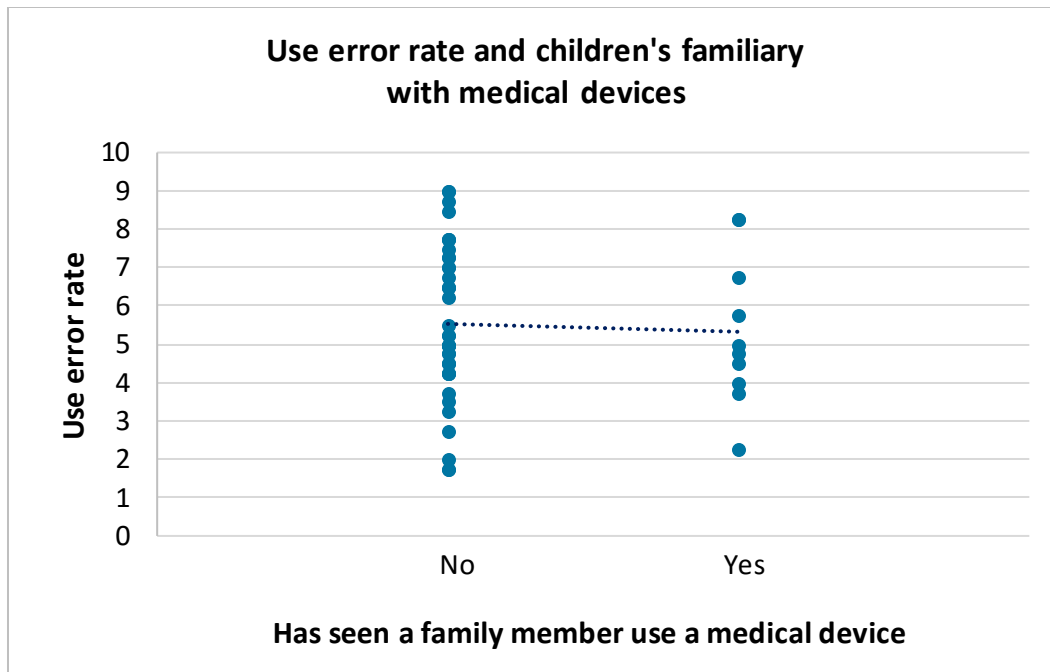
Figure 42: Use error rate and children's familiarity with medical devices

The results indicated that children's level of exposure to medical devices was not significantly associated with use error rate, *r* = -.043, *p* = .766. As shown in Table 6, none of the exploratory variables were significantly related to use error rate.

6.3.4. Summary of Correlation Results

All six developmental variables were statistically significantly correlated with the average number of use errors children committed with all devices (Table 7). However, the variables were strongly correlated with age (Table 5). Thus, the researcher reran the correlation matrix while controlling for the influence of age. Subsequently, six variables were no longer significant, including children's hand size, dominant and non-dominant grip strength, dominant and non-dominant dexterity, and functional numerosity score. Two variables remained significant

when controlling for age, including children's working memory scores and

reading level. Furthermore, use error rate was not significantly correlated with

children's gender or handedness, parent education level, or having seen a family

member use a medical device.

Table 7: Correlation between developmental variables and use error rate

Developmental variables	Correlation coefficient	Correlation coefficient
Control	None	Age
Hand length	-.696***	.023
Dominant grip strength	-.737***	-.054
Non-dominant grip strength	-.712***	-.070
Dominant dexterity	-.411**	-.079
Non-dominant dexterity	-.596***	-.075
Working memory capacity	-.685***	-.279*
Reading ability	-.751***	-.291*
Understanding of numerosity	-.793***	-.231

Note. * $p < .05$; ** $p < .01$; *** $p < .001$

The correlation results provided an estimation of the extent to which

each variable was related to committing use errors. However, a correlation

coefficient does not indicate whether variables are substantial predictors of the

outcome variable. Therefore, the researcher conducted a multiple regression to

determine whether age and the aforementioned variables predicted use error

rate.

6.3.5. Multiple Regression Analysis

The researcher ran a multiple regression to understand whether use error rate can be predicted based on children's age and developmental capabilities tested for during the primary study. The method was appropriate because several predictor variables were not independent of each another. When assessing whether one variable significantly predicts another, a multiple regression advantageously removes the influence of the subsequent independent variables. The procedure enabled the researcher to assess the unique contribution of each predictor variable. Notably, the researcher transformed age logarithmically.

The stepwise regression approach scanned for the predictor variable with the highest R^2 value, and then repeated the process until all remaining variables were non-significant. There were two observations with Cook's distance greater than one, adding undue effects on the model. Therefore, the researcher excluded those cases from the analysis. The multiple regression model statistically significantly predicted use error rate from children's age and working memory scores, $F(2,44) = 68.484$, $p < .001$. The R^2 for the overall model was 75.7% with an adjusted R^2 of 74.6%, a large size effect. Regression coefficients and standard errors can be found in Table 8.

Table 8: Summary of multiple regression analysis

Variable	*B*	*SE*$_B$	β
Intercept	21.836	1.570	-
Age	-14.555	2.114	-.700*
Working Memory	-.134	.060	-.227*

Note. * $p < .05$; *B* = unstandardized regression coefficient; SE_B = Standard error of the coefficient; β = standard coefficient

6.3.6. Discussion of User Characteristic-Related Use Errors

The researcher's initial prediction stated that children who were ten years of age and older would demonstrate task mastery with home health care devices. The findings suggest that children younger than ten were capable of performing medical device tasks safely and effectively. Nearly half of children who were nine years of age performed at a comparable level to adults. Importantly, children who were younger than nine did not demonstrate the ability to perform tasks with minimal error. The deviation from misuse at nine years of age has implications for (1) clinicians and parents assessing children's readiness to manage aspects of their own care and (2) medical device manufacturers who want to design their devices to meet children's needs effectively.

6.3.6.1. Implications for Parents and Clinicians

Children who were nine years of age and older had a greater opportunity for success. As such, parents and clinicians working with a child who is at least nine years of age can begin to assess the child's readiness to assume health-

related responsibilities. The aspects of biological and cognitive development that were significantly related to committing relatively few use errors included hand length, dexterity, grip strength, working memory capacity, reading ability, and understanding of numerosity. Importantly, two variables remained significant when controlling for the effect of age. The first variable pertained to children's working memory capacity. Scores on the working memory test were significantly negatively correlated with use error rate. Children who scored higher were less likely to commit use errors. Moreover, nearly half of children who performed the medical device tasks with minimal error were capable of recalling and sequencing at least four to five items, depending on the number of item categories.

Children's reading ability was the second factor that was significantly related to use error rate. Children who read at a higher grade level were less likely to commit use errors. Almost half of children who completed the medical device tasks with minimal error read above a fourth-grade level. However, the multiple regression results revealed that children's reading ability was not a significant predictor of use error rate.

Age and working memory accounted for a significant amount of the variance in use error rate. Therefore, parents and clinicians can optimize their ability to gauge readiness and provide more developmentally appropriate care by bearing in mind the child's age and working memory capacity. To determine with reasonable certainty whether children are fit to manage aspects of their own

care, parents and clinicians should ensure the child is at least nine years of age and capable of storing, processing, and manipulating at least four pieces of information. Parents and clinicians can assess children's working memory capacity rather quickly by showing images/reading the names of four items (e.g., foods, animals, appliances, clothing) consecutively, and then instructing the child to recite the objects verbally in some predetermined order (e.g., size order, alphabetical order).

Notably, parents and clinicians should take caution when generalizing these results to specific device types. Although the study results indicated that children's age and working memory capacity accounted for the greatest amount of variance in overall use error rate, there are additional factors to consider. To definitively determine whether a child is capable of self-administering treatment with a particular device, parents and clinicians need to have a comprehensive understanding of the device-specific biological and cognitive demands. Accordingly, a diabetes nurse educator or respiratory therapist could develop checklists to support clinical decision-making when determining whether a child is capable of using certain devices. An example of a checklist item for a blood glucose meter might include verifying that the child can solve problems involving subtraction with three-digit numbers before considering whether to allow him or her to interpret his or her own glucose test results.

6.3.6.2. Implications for Medical Device Manufacturers

The correlation and regression findings provided insight into the noticeable deviation in performance around nine years of age. Medical device manufacturers should consider drawing on the study results to inform the development of home health care device user interfaces and associated instructions and training. This section offers several design recommendations to mitigate risk associated with devices intended for use by children.

Recommendation 1. Ensure device user interfaces, IFUs, and training materials do not impose excessive demand on children's working memory. The device should not require children to process greater than four information elements simultaneously. For example, limit the number of menu options available on a software interface screen or modify scaffolding strategies during training such that the trainer models four task steps, and then provides the child with an opportunity to demonstrate comprehension and ask questions before proceeding.

Recommendation 2. Consult grade-level mathematic standards (e.g., the Common Core State Standards) to identify whether a typical nine-year-old child possesses the knowledge to demonstrate essential performance with the medical device in development. If not, consider implementing some form of risk mitigation. For example, physicians commonly prescribe nebulizer albuterol dosages between 2.5 and 5 ml to children diagnosed with asthma. However, a nine-year-old who lives in a state that has adopted the Common Core Standards

has not been introduced to the concept of decimal notation. Therefore, a device

manufacturer developing a nebulizer intended for use by children could increase

the likelihood that children measure medication doses accurately by color-coding

the medicine cup's graduation lines. Then, task success does not rely on the

child's understanding of numerosity.

Recommendation 3. The 5th-percentile nine-year-old is capable of

producing approximately 18 pounds of force with his or her hands (Reuben et al.,

2013). Therefore, consider designing devices such that the devices' components

require a maximum force of approximately 18 pounds. Doing so will ensure 95%

of nine-year-old children who desire to participate in their own care can perform

task steps that require grasping and applying force to an object, such as removing

a cap or pressing tabs to release a device component.

Recommendation 4. Ensure all device labeling is written at a third-grade

reading level or lower to increase the likelihood that an average nine-year-old

can read and comprehend the text. Device manufacturers might not find this

recommendation ideal when the medical device is intended for use by both the

pediatric and adult population. As such, a manufacturer could design the device

as intended for adults and then implement supplementary, "child-friendly"

design elements. For instance, include an additional IFU that is written at an

appropriate level for children. Device with software user interfaces could include

a child-driven design setting that simplifies the on-screen text and provides

additional learning aids, such as an animated graphic depicting a person inserting

the test strip into the meter in the correct orientation.

Accounting for children's biological and cognitive development during the

design process can potentially reduce the barriers to proper chronic disease

management for children.

6.3.7. Device Complexity and Use Error Rate

The primary study's second objective pertained to relating device

complexity to the prevalence of use errors. The researcher hypothesized that

there would be a positive relationship between device complexity and use error

rate. The researcher quantified device complexity using the Device-User

Complexity scoring procedure.

To assess the relationship between DUC score and use errors frequency,

the researcher ran a linear mixed-effects model (LMM) with repeated measures.

A correlation was not the appropriate statistical analysis because the primary

study's design had both within- and between-subject elements. Each child

(between-subjects) performed one task with each device (within-subjects). The

same 51 participants committed use errors with a device at each complexity

level. Unlike correlation analyses, which assume that all data points are

independent, LMM enabled the researcher to account for the variance in

intercepts across participants.

The relationship between device complexity (DUC score) and use error

rate showed significant variance in intercepts across participants $Var(\mu_{0j})$ = 3.238,

x^2 = .734, p < .001. Overall, DUC score statistically significantly predicted use

error rate, $F(1, 80.223)$ = 155.679, p < .001. As hypothesized, there was a positive

relationship between the variables. Children's use error rate increased by .052

use errors for each one-unit increase in DUC score. However, a visual inspection

of the scatter plot in Figure 43 alluded to a curvilinear relationship between the

variables. Subsequently, the researcher transformed DUC scores logarithmically

and performed an additional LMM analysis.
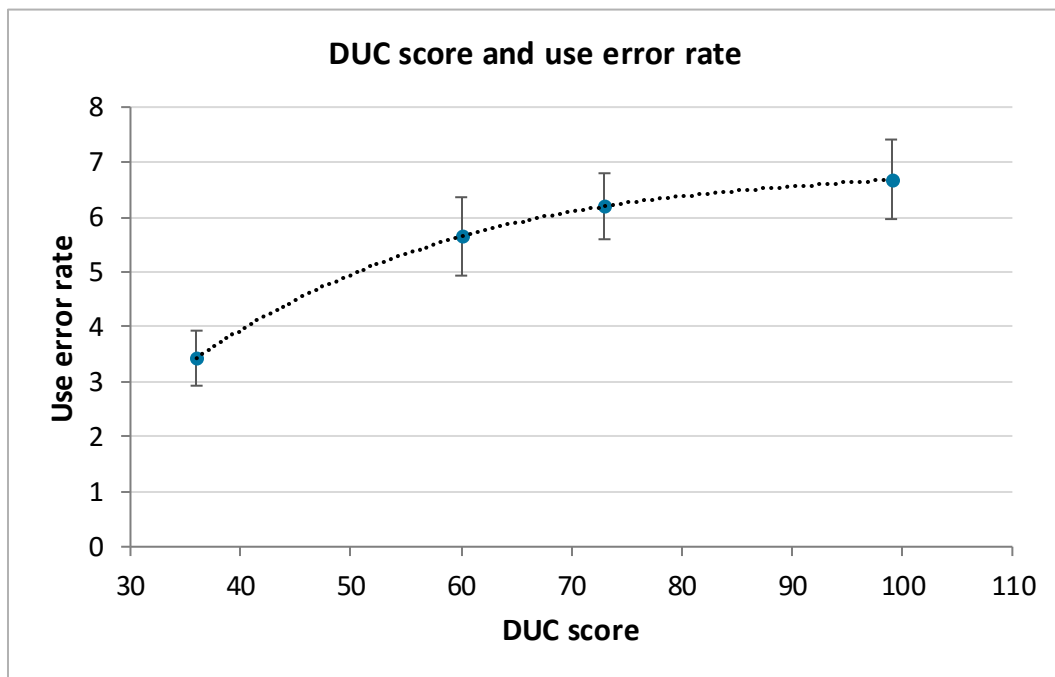
Figure 43: DUC score and average use error rate for child participants. Error bars represent the 95% CI for mean use error rate.

To determine whether use error rate and device complexity had a

curvilinear relationship, the researcher ran a second LMM analysis and compared

the models. The results indicated a decrease in log-likelihood (-2LL), confirming

that a curvilinear model is was a better fit than the initial linear model, x^2_{change} =

11.531, and yielded statistically significant results, $F(1, 76.913) = 200.510$, p < .001. Children's predicted use error rate is equal to $-8.397 + 7.697(\log DUC)$.

6.3.8. Discussion of Device Complexity-Related Use Errors

The results indicated that children committed significantly fewer use errors with devices that were relatively less complex, regardless of device type (i.e., blood glucose meters vs. nebulizers). Therefore, medical device manufacturers could use the DUC model to determine the extent to which the device will lead to use errors by children. A device with a relatively high DUC score delivers more comprehensive functionality and as such, likely requires a more exhaustive risk management process. Additionally, the metric affords comparisons among alternative functional designs. Adding a device function that significantly increases use error rate likely results in a greater quantity of hazards the manufacturer needs to address.

Lastly, the manufacturers can use the metric to facilitate decision-making when planning usability tests. Medical device manufacturers developing a product will conduct validation usability tests to provide the FDA with concrete evidence that the product presents minimal risk to its intended users in the representative use environment. As mentioned previously, the primary study results suggested that devices with high DUC scores increase the opportunity for error. Therefore, a medical device manufacturer should consider recruiting a larger sample of children when the DUC score is relatively high to increase the

likelihood of exposing the wide range of use errors that could occur when the device is placed in the real-world.

The following section discusses the model that explained the most variance in use error rate when accounting for both the device- and user-specific variables assessed during the primary study.

6.3.9. Final Regression Model

The researcher ran two additional LMM analyses to identify the model that accounted for the most variance in use error rate. The first model included the DUC score and children's age. The previous analyses identified both variables as significant predictors of use error rate. When entered into the same model, DUC score, $F(1, 73.879) = 13.951$, $p < .001$, and age, $F(1,49.860) = 95.440$, $p < .001$, remained significant predictors of use error rate. The -2LL value indicated that the model was a better fit than DUC score, $x^2_{change} = 53.374$, or age, $x^2_{change} = 124.653$, alone.

The stepwise multiple regression had identified working memory as a significant predictor of use error rate. Therefore, the second LMM the researcher conducted included DUC score, age, and working memory score as predictors. The model statistically significantly predicted use error rate and with higher accuracy than the model that omitted working memory score, $x^2_{change} = 21.565$. DUC score, $F(1, 72.012) = 194.220$, $p < .001$, age, $F(1, 50.382) = 27.394$, $p < .001$, and working memory, $F(1, 50.477) = 6.114$, $p = .017$, all significantly predicted the

number of use errors children committed with each medical device. Therefore,

children's predicted use error rate is equal to $5.774 + 7.591(\log DUC) - 11.833(\log Age) - .157(Working\ Memory)$. Table 9 specifies the final

regression model parameters.

Table 9: Final regression model parameters

Variable	b	SE$_b$	95% CI
Intercept	5.773	1.911	1.974, 9.573
Log(DUC)	7.591	.545	6.505, 8.677
Log(Age)	-11.833	2.261	-16.373, -7.293
Working Memory	-.157	.064	-.285, -.030

Note. b = Estimate; SE$_b$ = Standard error of the estimate

6.3.10. Discussion of Final Regression Model

The final regression model generated a fairly accurate prediction of use

error rate based on children's age and working memory score and DUC score.

The tool provides an independent metric for determining the rate of potential for

use errors by children when using home health care devices.

If parents and clinicians are going to relinquish health-related

responsibilities to children, they need to feel confident that the child is capable

of performing with minimal error. This metric has the capability to estimate the

extent to which a medical device presents a safety risk to children regarding the

administration of their own care. As such, medical device manufacturers could

employ the DUC model to determine a device's complexity level, and then utilize

the final regression model to rate the potential for use errors by children,
thereby simplifying and streamlining any subsequent safety testing.

6.4. Study Limitations

The primary study had some limitations. The study included a relatively small sample size of adult participant. Recruiting a larger sample of adults would have afforded a more accurate estimation of the maximum number of use errors the adult population is likely to commit when using nebulizers and blood glucose meters. This would have increased the likelihood of correctly identifying the children who performed at a comparable level to adults and were likely capable of performing tasks with home health care devices with minimal error.

In addition, the participants the researcher recruited had not been diagnosed with diabetes and most had not received an asthma diagnosis. As such, the results do not account for the prevalence of developmental deficits in children who have these conditions. For example, children diagnosed with type 1 diabetes are prone to reduced cognitive functioning, potentially increasing use error risk. Moreover, the results to do take into account that children diagnosed with diabetes or asthma tend to experience symptoms that might impact their ability to perform device task effectively. For instance, a diabetic child experiencing symptoms of hyper or hypoglycemia is likely to be in a very emotional state, which might influence decision-making. Therefore, parents, clinicians, and medical device manufacturers should bear in mind the

developmental factors and symptoms associated with given chronic diseases, and their potential effect on performance before utilizing the primary study data.

6.5. Future Research

There are several opportunities for future research. The researcher provided recommendations for developing device user interfaces and associated materials to decrease use error rate among children. Supplemental research might involve developing several prototypes that implement the design suggestions, and then conducting a usability study to determine the extent to which the modifications mitigate the risks.

In addition, the primary study tested for several aspects of child development. However, there were domains that were not included in the study, which might influence use error rate, such as children's emotional and episodic memory development. As children grow, their emotional reactions to environmental changes evolve. They develop the emotional competence to manage feelings of stress, frustration, fear, and sadness effectively. Performing a complex cognitive task such as working with a home health care device might induce negative emotions. Thus, a valuable follow up study might involve administering functional assessments to measure children's capacity to regulate their emotions and examine the magnitude of its relation to use error frequency and distinct use error types.

Episodic memory involves recalling details of personal experiences and events that occurred at a particular time and place in the past, as well as the associated emotions. As such, the extent of children's episodic memory development might influence their ability to learn from their mistakes. A future study might involve modifying the aforementioned primary study such that each child performs the medical device tasks and then, after a decay period, returns for a second session to repeat the tasks. The study could include a metric to assess children's episodic memories (e.g., NIH Toolbox Picture Sequence Memory Test (Tulsky et al., 2013)) to determine the age when children become capable of modifying their behavior based on incident that occurred (e.g., learn to perform the task correctly), and potentially predict this deviation based on the child's episodic memory test score.

7. Conclusion

There is a scarcity of literature indicating the age that children become capable of using home health care devices safely and effectively and the developmental variables associated with success. Nine years of age marked a substantial rise in the number of children who performed nebulizer and blood glucose meter tasks with minimal error. Furthermore, children's age and working memory capacity accounted for a significant proportion of the variance in use error rate. Child participants who were younger than nine or had a cognitive capacity of fewer than four items appeared to commit an excessive number of use errors relative to adult participants.

Furthermore, this is the first study known to relate device complexity to the prevalence of use errors. Children were significantly less likely to commit use errors when device complexity was relatively low. Subsequently, accounting for DUC score and children's age and working memory capacity showed to explain the most variance in use error rate. The model significantly predicted the rate of potential for use errors by children when using home health care devices, thereby providing a method for medical device manufacturers to systematically determine the extent to which their product will lead to use errors. Moreover, this tool can increase the confidence in decision-making by parents, teachers, school nurses, pediatric educators who decide the type of device to recommend to a child, among others.

8. References

Ameri, F., Summers, J. D., Mocko, G. M., & Porter, M. (2008). Engineering design complexity: an investigation of methods and measures. *Research in Engineering Design*, *19*(2-3), 161–179. http://doi.org/10.1007/s00163-008-0053-2

Anderson, J. R., & Jeffries, R. (1985). Novice LISP Errors: Undetected Losses of Information from Working Memory. *Human–Computer Interaction*, *1*(2), 107–131. http://doi.org/10.1207/s15327051hci0102_2

Association for the Advancement of Medical Instrumentation. (2009). ANSI/AAMI HE75: 2009. *Human Factors engineering–Design of Medical Devices*.

Bashir, H. A., & Thomson, V. (1999). Estimating Design Complexity. *Journal of Engineering Design*, *10*(3), 247–257. http://doi.org/10.1080/095448299261317

Bredekamp, S., & National Association for the Education of Young Children (Eds.). (1987). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8* (Expanded ed). Washington, DC: National Association for the Education of Young Children.

Bull, R., & Scerif, G. (2001). Executive Functioning as a Predictor of Children's Mathematics Ability: Inhibition, Switching, and Working Memory. *Developmental Neuropsychology*, *19*(3), 273–293. http://doi.org/10.1207/S15326942DN1903_3

Center for Devices and Radiological Health, U.S. Food and Drug Administration. (n.d.). Products and Medical Procedures - Pediatric Medical Devices. Retrieved July 17, 2016, from http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/ucm135104.htm

Centers for Disease Control and Prevention. (2014). *Summary Health Statistics Tables for U.S. Children: National Health Interview Survey, 2014, Table C-1* (Summary Health Statistics: National Health Interview Survey).

Claasen, T. A. C. M. (2003). Platform design: the next paradigm shift to deal with complexity (pp. 8–12). IEEE. http://doi.org/10.1109/VTSA.2003.1252538

Clarke, W. L. (2011). Behavioral Challenges in the Management of Childhood Diabetes. *Journal of Diabetes Science and Technology*, *5*(2), 225–228. http://doi.org/10.1177/193229681100500203

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, N.J: L. Erlbaum Associates.

Crespo-Varela, J. R., Medina, L. A., & Kremer, G. E. O. (Eds.). (2012). An Analysis of the Relationship of Medical Device Complexity and FDA. In *Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Chicago, Illinois, USA: American Society of Mechanical Engineers.

Erikson, E. H. (1956). The Problem of Ego Identity. *Journal of the American Psychoanalytic Association*, *4*(1), 56–121. http://doi.org/10.1177/000306515600400104

Gogtay, N., & Thompson, P. M. (2010). Mapping gray matter development: Implications for typical development and vulnerability to psychopathology. *Brain and Cognition*, *72*(1), 6–15. http://doi.org/10.1016/j.bandc.2009.08.009

Hitch, G. J., Towse, J. N., & Hutton, U. (2001). What limits children's working memory span? Theoretical accounts and applications for scholastic development. *Journal of Experimental Psychology: General*, *130*(2), 184–198. http://doi.org/10.1037/0096-3445.130.2.184

Hodes, R. J., Insel, T. R., Landis, S. C., & On behalf of the NIH Blueprint for Neuroscience Research. (2013). The NIH Toolbox: Setting a standard for biomedical research. *Neurology*, *80*(Issue 11, Supplement 3), S1–S1. http://doi.org/10.1212/WNL.0b013e3182872e90

Howe, T. R., Trotter, J. S., Davis, A. S., Schofield, J. W., Allen, L., Millians, M., … McLellan, T. (2011). Adaptive Behavior. In S. Goldstein & J. A. Naglieri (Eds.), *Encyclopedia of Child Behavior and Development* (pp. 31–32). Boston, MA: Springer US. Retrieved from http://www.springerlink.com/index/10.1007/978-0-387-79061-9_52

International Diabetes Federation. (2015). *IDF Diabetes Atlas: Seventh Edition 2015*.

International Electrotechnical Committee (IEC). (2015). *IEC 62366-1:2015: Medical Devices-Part 1: Application of Usability Engineering to Medical Devices.* Geneva, Switzerland.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, *99*(1), 122–149.

La Pray, M., & Ross, R. (1969). The graded word list: Quick gauge of reading ability. *Journal of Reading*, *12*(4), 305–307.

Lenroot, R. K., & Giedd, J. N. (2006). Brain development in children and adolescents: Insights from anatomical magnetic resonance imaging. *Neuroscience & Biobehavioral Reviews*, *30*(6), 718–729. http://doi.org/10.1016/j.neubiorev.2006.06.001

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30–46. http://doi.org/10.1037/1082-989X.1.1.30

McNally, K., Rohan, J., Pendley, J. S., Delamater, A., & Drotar, D. (2010). Executive Functioning, Treatment Adherence, and Glycemic Control in Children With Type 1 Diabetes. *Diabetes Care*, *33*(6), 1159–1162. http://doi.org/10.2337/dc09-2116

Miller, G. A. (1994). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *101*(2), 343–352. http://doi.org/10.1037//0033-295X.101.2.343

Munakata, Y., Casey, B. J., & Diamond, A. (2004). Developmental cognitive neuroscience: progress and potential. *Trends in Cognitive Sciences*, *8*(3), 122–128. http://doi.org/10.1016/j.tics.2004.01.005

Myatchin, I., & Lagae, L. (2013). Developmental changes in visuo-spatial working memory in normally developing children: event-related potentials study. *Brain & Development*, *35*(9), 853–864. http://doi.org/10.1016/j.braindev.2012.11.005

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). Common Core State Standards in Mathematics. National Governors Association Center for Best Practices, Council of Chief State School Officers, Washington D.C.

Naughten, E. (1982). At What Age Do Diabetic Children Give Their Own Injections? *Archives of Pediatrics & Adolescent Medicine*, *136*(8), 690. http://doi.org/10.1001/archpedi.1982.03970440034009

Norman, D. A. (2002). *The design of everyday things* (1st Basic paperback). New York: Basic Books.

Perwien, A. R., Johnson, S. B., Dymtrow, D., & Silverstein, J. (2000). Blood glucose monitoring skills in children with Type I diabetes. *Clinical Pediatrics*, *39*(6), 351–357.

Piaget, J., Inhelder, B., & Weaver, H. (1969). *The Psychology of the child* (Nachdr). New York: Basic Books, Inc.

Reuben, D. B., Magasi, S., McCreath, H. E., Bohannon, R. W., Wang, Y.-C., Bubela, D. J., … Gershon, R. C. (2013). Motor assessment using the NIH Toolbox. *Neurology*, *80*(Issue 11, Supplement 3), S65–S75. http://doi.org/10.1212/WNL.0b013e3182872e01

Riggs, K. J., McTaggart, J., Simpson, A., & Freeman, R. P. J. (2006). Changes in the capacity of visual working memory in 5- to 10-year-olds. *Journal of Experimental Child Psychology*, *95*(1), 18–26. http://doi.org/10.1016/j.jecp.2006.03.009

Roy, R., Evans, R., Low, M. J., & Williams, D. K. (2011). Addressing the impact of high levels of product variety on complexity in design and manufacture. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *225*(10), 1939–1950. http://doi.org/10.1177/0954405411407670

Santrock, J. W. (2011). *Life-span development* (13th ed). New York, NY: McGraw-Hill Higher Education.

Santrock, J. W. (2013). *Child development* (Fourteenth edition). Dallas: University of Texas.

Scott, L. K. (2013). Developmental Mastery of Diabetes-Related Tasks in Children. *Nursing Clinics of North America*, *48*(2), 329–342. http://doi.org/10.1016/j.cnur.2013.01.015

Smith Jr, C. F., & Harrison, M. D. (1983). A Comparative Analysis of Three Widely Used Graded Word Reading Tests.

Sohn, Y. W., & Doane, S. M. (2003). Roles of working memory capacity and long-term working memory skill in complex task performance. *Memory & Cognition*, *31*(3), 458–466. http://doi.org/10.3758/BF03194403

Sroufe, L. A., Cooper, R. G., DeHart, G., & Bronfenbrenner, U. (1992). *Child development: its nature and course* (2nd ed). New York: McGraw-Hill.

Tulsky, D. S., Carlozzi, N. E., Chevalier, N., Espy, K. A., Beaumont, J. L., & Mungas, D. (2013). V. NIH TOOLBOX COGNITION BATTERY (CB): MEASURING WORKING MEMORY: NIH TOOLBOX COGNITION BATTERY (CB). *Monographs of the Society for Research in Child Development*, *78*(4), 70–87. http://doi.org/10.1111/mono.12035

U.S. Department of Health and Human Services, Food and Drug Administration. (2001). *Guidance on Medical Device Patient Labeling*. Retrieved from http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm070782.htm

U.S. Department of Health and Human Services, Food and Drug Administration. (2016). *Applying Human Factors and Usability Engineering to Medical Devices*.

Wang, Y.-C., Bohannon, R. W., Kapellusch, J., Garg, A., & Gershon, R. C. (2015). Dexterity as measured with the 9-Hole Peg Test (9-HPT) across the age span. *Journal of Hand Therapy*, *28*(1), 53–60. http://doi.org/10.1016/j.jht.2014.09.002

Weinger, M. B., Gardner-Bonneau, D., Wiklund, M. E., & Kelly, L. M. (Eds.). (2011). *Handbook of human factors in medical device design*. Boca Raton, FL: CRC Press.

Wiklund, M. E., Kendler, J., & Strochlic, A. Y. (2016). *Usability testing of medical devices*. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1132075

Zhang, H., Li, Y.-F., & Tan, H. B. K. (2010). Measuring design complexity of semantic web ontologies. *Journal of Systems and Software*, *83*(5), 803–814. http://doi.org/10.1016/j.jss.2009.11.735

9. Appendices

Appendix A. Preliminary Study DUC Functional Trees

Kitchen Thermometer 1: DUC 8
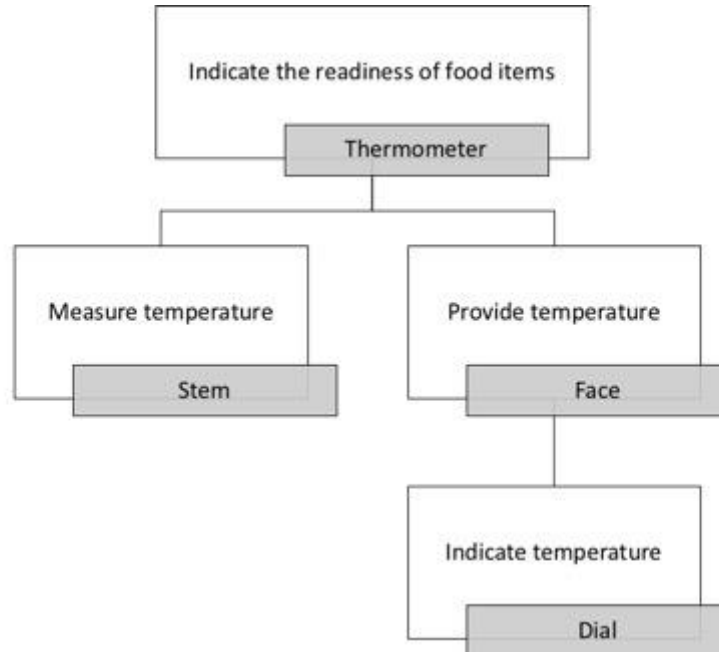
Wrist Watch 1: DUC 14

Coffee Maker 1: DUC 29

A hierarchical diagram showing the functional decomposition of a Coffee Maker.

- Provide brewed coffee — Coffee Maker
  - Hold brewed coffee — Decanter
    - Deactivate coffee brewing — On/Off Button
  - Filter coffee grounds — Coffee Filter
    - Maintain coffee temperature — Warming Plate
  - Brew coffee — Coffee Maker Unit
    - Initiate coffee brewing — On/Off Button
    - Hold coffee grounds — Filter Basket
    - Enable monitoring of water level — Reservoir
      - Enable monitoring of water level — Water Window
    - Hold water — Reservoir

Kitchen Thermometer 2: DUC 44

Watch 2: DUC 66
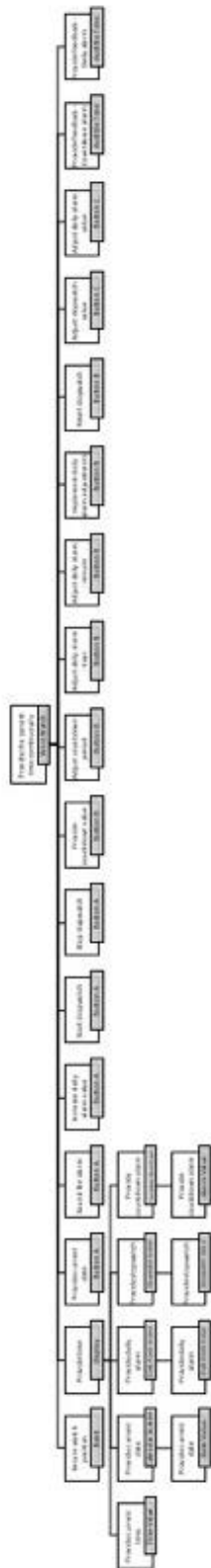
Coffee Maker 2: DUC 112
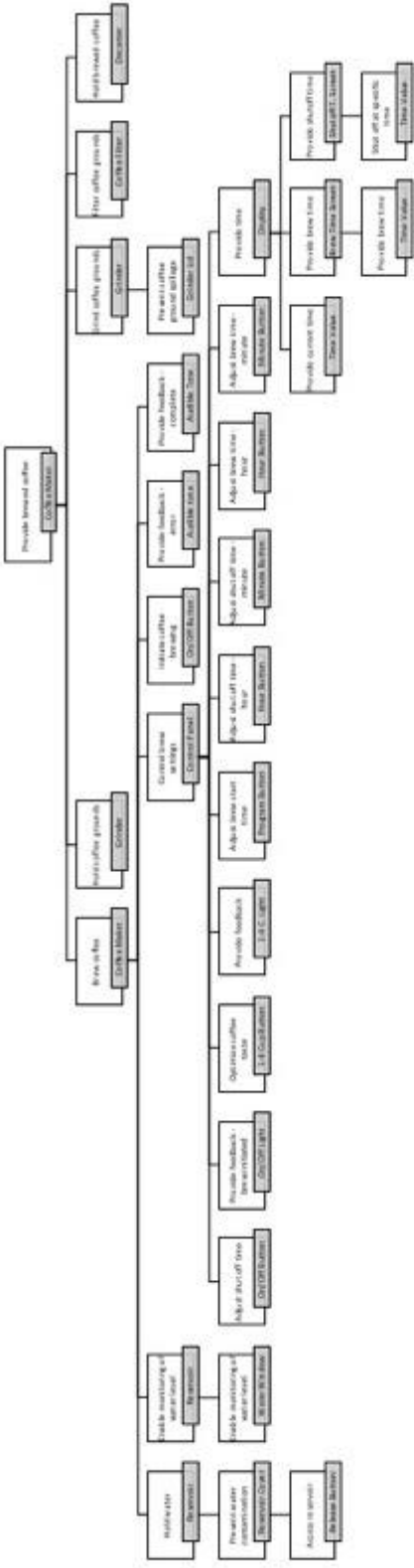
Appendix B. Preliminary Study SPSS Output

Interrater Reliability

Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.926[a]	.814	.987	125.413	5	55	.000
Average Measures	.993	.981	.999	125.413	5	55	.000

Curve Estimation Regression Analysis

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
.991	.983	.979	1.628

Appendix C. Primary Study DUC Functional Trees

Nebulizer 1: DUC 36
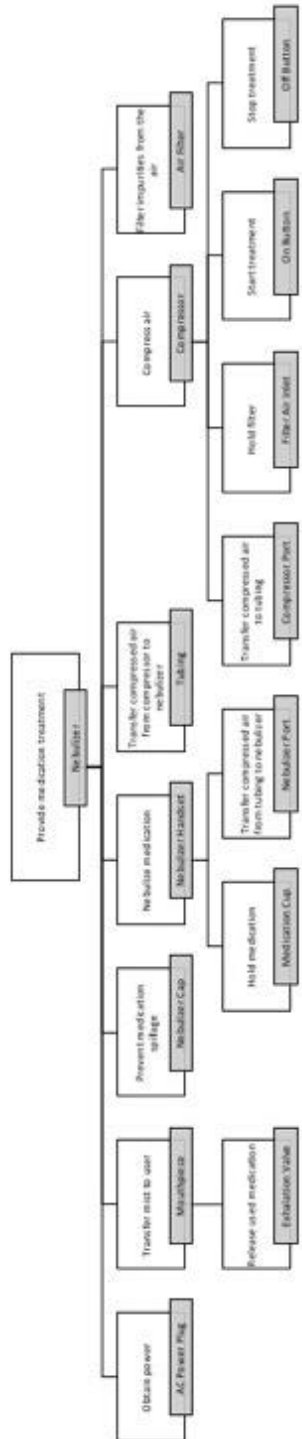
Blood Glucose Meter 1: DUC 60
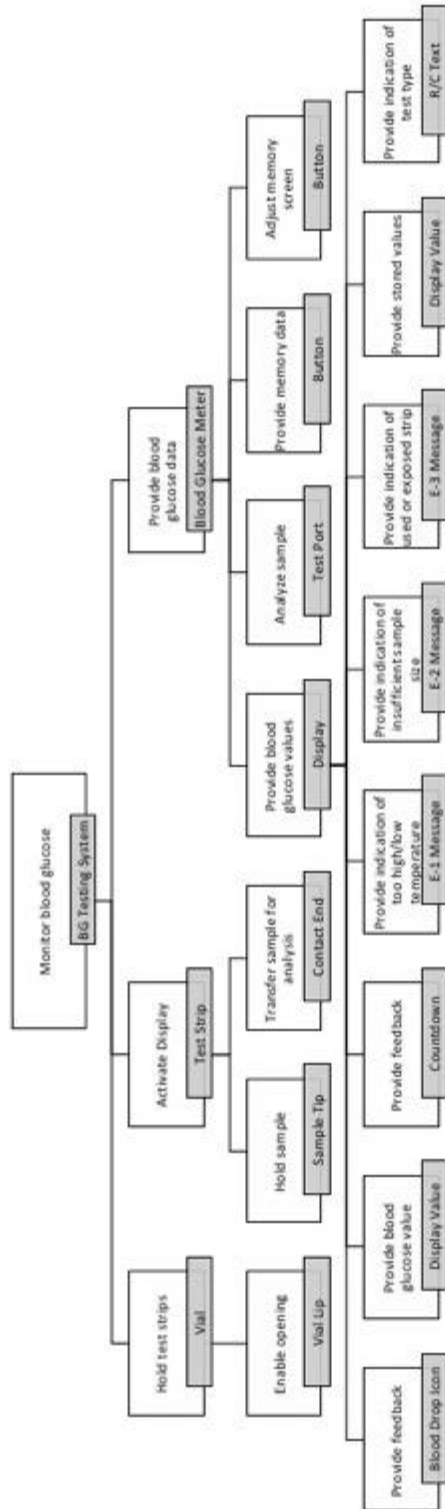
Nebulizer 2: DUC 73
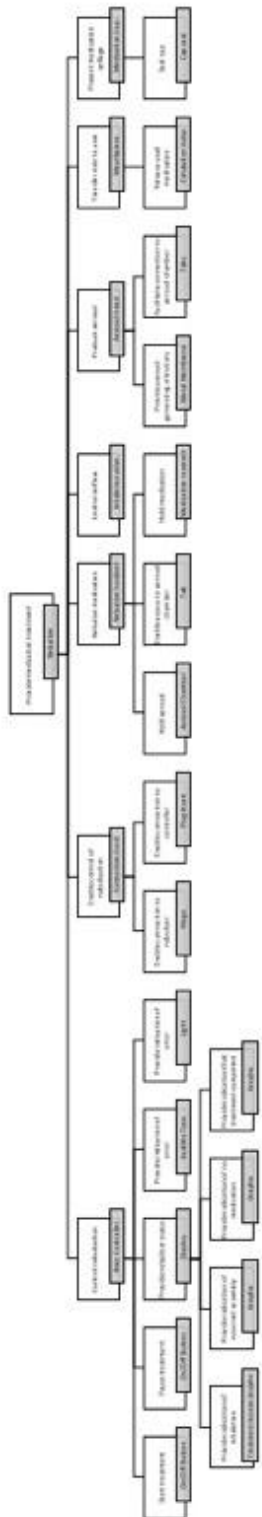
Blood Glucose Meter 2: DUC 99
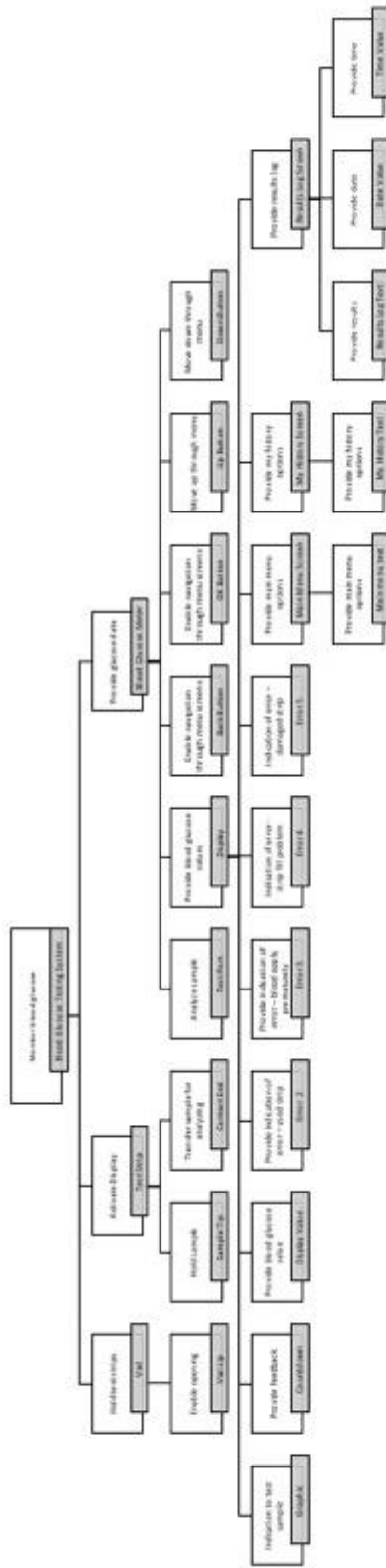
Appendix D. Primary Study SPSS Output

Curve Estimation Regression – Age and Use Error Rate

R	R Square	Adjusted R Square	Std. Error of the Estimate
.835	.697	.691	1.070

The independent variable is Age.

Pearson Correlation – Age and Use Error Rate

		Average Use Errors - All Devices	Age
Average Use Errors - All Devices	Pearson Correlation	1	-.835**
	Sig. (2-tailed)		.000
	N	51	51
Age	Pearson Correlation	-.835**	1
	Sig. (2-tailed)	.000	
	N	51	51

Partial Correlation – Biological Variables and Use Error Rate Controlling for Age

Control Variables			Average Use Errors - All Devices	Age
-none-[a]	Average Use Errors – All Devices	Correlation	1.000	-.829
		Significance (1-tailed)	.	.000
		df	0	49
	Hand Length	Correlation	-.696	.848
		Significance (1-tailed)	.000	.000
		df	49	49
	Dominant Dexterity (Seconds)	Correlation	.411	-.541
		Significance (1-tailed)	.001	.000
		df	49	49
	Non-Dominant Dexterity (Seconds)	Correlation	.596	-.682
		Significance (1-tailed)	.000	.000
		df	49	49
	Dominant Grip Strength (Pounds)	Correlation	-.737	.871
		Significance (1-tailed)	.000	.000
		df	48	48
	Non-Dominant Grip Strength (Pounds)	Correlation	-.712	.832
		Significance (1-tailed)	.000	.000
		df	48	48

			Average Use Errors	
	Age	Correlation	-.829	1.000
		Significance (1-tailed)	.000	.
		df	49	0
Age	Average Use Errors – All Devices	Correlation	1.000	
		Significance (1-tailed)	.	
		df	0	
	Hand Length	Correlation	.023	
		Significance (1-tailed)	.436	
		df	48	
	Dominant Dexterity (Seconds)	Correlation	-.079	
		Significance (1-tailed)	.292	
		df	48	
	Non-Dominant Dexterity (Seconds)	Correlation	.075	
		Significance (1-tailed)	.303	
		df	48	
	Dominant Grip Strength (Pounds)	Correlation	-.054	
		Significance (1-tailed)	.356	
		df	47	
	Non-Dominant Grip Strength (Pounds)	Correlation	-.070	
		Significance (1-tailed)	.316	
		df	47	

Partial Correlation – Cognitive Variables and Use Error Rate Controlling for Age

Control Variables			Average Use Errors - All Devices	Age
-none-[a]	Average Use Errors - All Devices	Correlation	1.000	-.829
		Significance (1-tailed)	.	.000
		df	0	49
	Working Memory - No. Correct	Correlation	-.685	.690
		Significance (1-tailed)	.000	.000
		df	48	48
	Reading Ability - Grade Level	Correlation	-.751	.785
		Significance (1-tailed)	.000	.000
		df	49	49
	Numerosity - No. Correct	Correlation	-.793	.884
		Significance (1-tailed)	.000	.000
		df	47	47
	Age	Correlation	-.829	1.000
		Significance (1-tailed)	.000	.
		df	49	0
Age	Average Use Errors - All Devices	Correlation	1.000	
		Significance (1-tailed)	.	
		df	0	

Working Memory - No. Correct	Correlation	-.279
	Significance (1-tailed)	.026
	df	47
Reading Ability - Grade Level	Correlation	-.291
	Significance (1-tailed)	.020
	df	48
Numerosity - No. Correct	Correlation	-.231
	Significance (1-tailed)	.057
	df	46

			Average Use Errors - All Devices
Spearman's rho	Average Use Errors - All Devices	Correlation Coefficient	1.000
		Sig. (1-tailed)	.
	Functional Numerosity	Correlation Coefficient	-.795**
		Sig. (1-tailed)	.000

Pearson Correlation – Raw Working Memory Score and Modified Score

		Working Memory - Raw Score	Working Memory - Modified Score
Working Memory - Raw Score	Pearson Correlation	1	.976**
	Sig. (2-tailed)		.000
	N	50	50
Working Memory - Modified Score	Pearson Correlation	.976**	1
	Sig. (2-tailed)	.000	
	N	50	51

**. Correlation is significant at the 0.01 level (2-tailed).

Partial Correlation – Reading the BG Result Incorrectly and Numerosity Score Controlling for Age

Control Variables			Numerosity Score	Read BG Results Incorrectly	Age
-none-[a]	Numerosity Score	Correlation	1.000	-.671	.884
		Significance (2-tailed)	.	.000	.000
		df	0	47	47
	Read BG Results Incorrectly	Correlation	-.671	1.000	-.641
		Significance (2-tailed)	.000	.	.000
		df	47	0	49
	Age	Correlation	.884	-.641	1.000

		Significance (2-tailed)	.000	.000	.
		df	47	49	0
Age	Numerosity Score	Correlation	1.000	-.293	
		Significance (2-tailed)	.	.044	
		df	0	46	
	Read BG results incorrectly	Correlation	-.293	1.000	
		Significance (2-tailed)	.044	.	
		df	46	0	

a. Cells contain zero-order (Pearson) correlations.

Pearson Correlation – Exploratory Variables and Use Error Rate

		Average Use Errors - All Devices
Handedness	Pearson Correlation	-.076
	Sig. (2-tailed)	.595
	N	51
Gender	Pearson Correlation	-.021
	Sig. (2-tailed)	.885
	N	51
Parent's Highest Level of Education Completed	Pearson Correlation	-.125
	Sig. (2-tailed)	.383
	N	51
Family Member Device Use	Pearson Correlation	-.043
	Sig. (2-tailed)	.766
	N	51

Multiple Regression Analysis

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.854[a]	.729	.723	1.0268	.729	121.275	1	45	.000	
2	.870[b]	.757	.746	.9842	.028	4.977	1	44	.031	2.386

a. Predictors: (Constant), Age
b. Predictors: (Constant), Age, Working Memory Score
c. Dependent Variable: Average Use Errors - All Devices

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	127.855	1	127.855	121.275	.000[b]
	Residual	47.442	45	1.054		
	Total	175.297	46			
2	Regression	132.676	2	66.338	68.484	.000[c]
	Residual	42.621	44	.969		
	Total	175.297	46			

a. Dependent Variable: Average Use Errors - All Devices
b. Predictors: (Constant), Age
c. Predictors: (Constant), Age, Working Memory Score

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B Lower Bound	Upper Bound	Correlations Zero-order	Partial	Part	Collinearity Statistics Toler.	VIF
1	(Constant)	22.767	1.579		14.421	.000	19.587	25.947					
	Age	-17.769	1.614	-.854	-11.012	.000	-21.019	-14.519	-.854	-.854	-.854	1.000	1.000
2	(Constant)	21.836	1.570		13.910	.000	18.672	24.999					
	Age	-14.555	2.114	-.700	-6.886	.000	-18.815	-10.295	-.854	-.720	-.512	.535	1.868
	Working Memory Score	-.134	.060	-.227	-2.231	.031	-.255	-.013	-.703	-.319	-.166	.535	1.868

a. Dependent Variable: Average Use Errors - All Devices

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics Tolerance	VIF	Minimum Tolerance
1	Working Memory Score	-.227[b]	-2.231	.031	-.319	.535	1.868	.535
	Reading - Grade Level	-.205[b]	-1.642	.108	-.240	.372	2.686	.372
	Hand Length	.027[b]	.194	.847	.029	.309	3.240	.309
	Dominant Dexterity	-.031[b]	-.342	.734	-.052	.750	1.332	.750
	Non-Dominant Dexterity	.040[b]	.356	.724	.054	.496	2.018	.496
	Dominant Grip Strength	-.043[b]	-.292	.772	-.044	.281	3.560	.281
	Non-Dominant Grip Strength	-.088[b]	-.670	.506	-.101	.352	2.841	.352

	Beta In	t	Sig.	Partial Correlation	Tolerance	VIF	Minimum Tolerance
Numerosity	-.201[b]	-1.210	.233	-.179	.215	4.652	.215
2 Reading Ability - Grade Level	-.140[c]	-1.106	.275	-.166	.343	2.912	.329
Hand Length	-.026[c]	-.191	.849	-.029	.299	3.344	.212
Dominant Dexterity	-.002[c]	-.023	.982	-.004	.733	1.364	.418
Non-Dominant Dexterity	.026[c]	.241	.811	.037	.494	2.025	.360
Dominant Grip Strength	-.090[c]	-.629	.533	-.096	.275	3.635	.203
Non-Dominant Grip Strength	-.118[c]	-.933	.356	-.141	.348	2.871	.250
Numerosity Score	-.104[c]	-.614	.542	-.093	.196	5.106	.196

a. Dependent Variable: Average Use Errors - All Devices

b. Predictors in the Model: (Constant), Age

c. Predictors in the Model: (Constant), Age, Working Memory Score

Linear Mixed-Effects Model – Use Error Rate and DUC Score: Pre-Logarithmic Transformation

Information Criteria[a]

-2 Log Likelihood	762.936
Akaike's Information Criterion (AIC)	772.936
Hurvich and Tsai's Criterion (AICC)	773.250
Bozdogan's Criterion (CAIC)	794.352
Schwarz's Bayesian Criterion (BIC)	789.352

The information criteria are displayed in smaller-is-better form.

a. Dependent Variable: use errors.

Type III Tests of Fixed Effects[a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	120.456	26.017	.000
complexity	1	73.078	174.088	.000

a. Dependent Variable: use errors.

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	1.925943	.377589	120.456	5.101	.000	1.178373	2.673514
complexity	.052503	.003979	73.078	13.194	.000	.044572	.060433

a. Dependent Variable: use errors.

Linear Mixed-Effects Model – Use Error Rate and DUC Score: Post Logarithmic Transformation

Information Criteria[a]

-2 Log Likelihood	751.405
Akaike's Information Criterion (AIC)	761.405
Hurvich and Tsai's Criterion (AICC)	761.719
Bozdogan's Criterion (CAIC)	782.821
Schwarz's Bayesian Criterion (BIC)	777.821

The information criteria are displayed in smaller-is-better form.
a. Dependent Variable: use errors.

Type III Tests of Fixed Effects[a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	89.342	68.708	.000
Complexity-log	1	76.913	200.510	.000

a. Dependent Variable: use errors.

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept	-8.397369	1.013074	89.342	-8.289	.000	-10.410219	-6.384519
Complexity-log	7.697130	.543577	76.913	14.160	.000	6.614711	8.779550

a. Dependent Variable: use errors.

Linear Mixed-Effects Model – Use Error Rate with DUC Score and Age

Information Criteria[a]

-2 Log Likelihood	698.031
Akaike's Information Criterion (AIC)	710.031
Hurvich and Tsai's Criterion (AICC)	710.474
Bozdogan's Criterion (CAIC)	735.731
Schwarz's Bayesian Criterion (BIC)	729.731

The information criteria are displayed in smaller-is-better form.
a. Dependent Variable: use errors.

Type III Tests of Fixed Effects[a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	88.988	14.553	.000
Complexity-log	1	73.879	194.643	.000
Age-log	1	49.860	95.440	.000

a. Dependent Variable: use errors.

Estimates of Fixed Effects[a]

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	7.150112	1.874267	88.988	3.815	.000	3.425976	10.874249
Complexity -log	7.672291	.549928	73.879	13.951	.000	6.576506	8.768076
Age-log	-16.015309	1.639340	49.860	-9.769	.000	-19.308251	-12.722368

a. Dependent Variable: use errors.

Linear Mixed-Effects Model – Use Error Rate with DUC Score, Age, and Working Memory

Information Criteria[a]

-2 Log Likelihood	676.466
Akaike's Information Criterion (AIC)	690.466
Hurvich and Tsai's Criterion (AICC)	691.072
Bozdogan's Criterion (CAIC)	720.305
Schwarz's Bayesian Criterion (BIC)	713.305

The information criteria are displayed in smaller-is-better form.

a. Dependent Variable: use errors.

Type III Tests of Fixed Effects[a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	86.337	9.125	.003
Complexity-log	1	72.012	194.220	.000
Age-log	1	50.382	27.394	.000
Working memory	1	50.447	6.114	.017

a. Dependent Variable: use errors.

Estimates of Fixed Effects[a]

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	5.773838	1.911333	86.337	3.021	.003	1.974446	9.573230
Complexity-log	7.591371	.544719	72.012	13.936	.000	6.505496	8.677247
Age-log	-11.833134	2.260869	50.382	-5.234	.000	-16.373371	-7.292898
Working memory	-.157306	.063617	50.447	-2.473	.017	-.285056	-.029556

a. Dependent Variable: use errors.