Axiomatic design of a man-machine interface for Alzheimer’s patient care

Sergio A. Navarro-Tuch1,2, Jack Gammack1, David Kang1, Sang-Gook Kim1

1 Park Center for Complex Systems, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

2 School of Engineering and Sciences, Tecnológico de Monterrey, Mexico City, Mexico

E-mail: sangkim@mit.edu

Abstract. Axiomatic thinking framework is used to develop a novel concept of “Care-Toy,” which helps caregivers of Alzheimer’s and dementia patients by providing much needed respite for themselves without compromising patients’ safety. “Care-Toy” interacts with a patient via stimulating images and sounds for touch interactions, actively keeping the patient’s attention for a needed duration as well as monitoring the patient’s status or alerting the caregiver if necessary. Systems solution for this concept could be effectively elaborated with the Axiomatic Design framework, leading to the development of a key module intended for measuring and maintaining the attention of the patient during the respite of caregivers. Gaze detection is the key module based on deep learning-based facial recognition system to monitor whether “Care-Toy” attracts the patient’s interest and maintains attention, otherwise to alert the caregiver. The module-junction structure could be derived from the decomposed functional requirement (FR)-design parameter (DP) tree structure and design matrices of them, which effectively defines the software system structure for this product.

1. Introduction

As part of the current global demographic transition, we have seen a rise in the average age of the population in almost every part of the globe. The global population aged 60 years or older numbered 962 million in 2017, more than twice as large as in 1980, and is expected to double again by 2050 to reach nearly 2.1 billion [1]. Being that age is the strongest known factor to the onset of degenerative mental diseases such as dementia [2], [3], and since 60 to 70% of dementia causes are related to Alzheimer’s [4], it is very important to develop strategies and tools for the care of such individuals.

As the individuals’ disease progresses, the dependence towards the caregiver, mostly family members of the patients, increases, as does the hardship on the whole family. Severe symptoms can frequently hinder the caregivers’ capability to manage necessary daily tasks, such as food preparation, supply acquisition, and cleaning, as losing sight of the patient could result in the patient wandering off and getting hurt in such a time period. It is essential to provide a much needed respite for a caregiver so that minimum quality of living can be maintained.

In order to aid the caregivers and provide them with a kind of short respite, about 30 minutes to 1 hour, the idea of “Care-Toy” was conceived, which is similar to the concept of providing toys to toddlers in order to make a short respite for parents doing essential tasks. As a toy must attract toddlers’ interest for a duration of respite while not harming them at all, Care-Toy should be so. In addition to these basic
functions, Care-Toy can bring more functions for the goodness of patients enabled by the widely available smart tablets. Based on this initial concept, Axiomatic Design thinking framework has been applied [5] to identify and define the FRs and DPs needed to detail the design of the system. Once the key modules are determined from the upper-level FR-DP tree structure, the main focus of this development becomes the “Attention Module” which is further explored in the section following the general design concept development.

2. Conceptual Design

An overview of the highest level functional requirements (FR) and design parameters (DP) are first presented to address the needs of the customer. An outline and brief discussion of the full Axiomatic decomposition of the system follows. Lastly, some of the most important modules are expanded upon and implemented.

2.1. High-Level System Design

Solution development requires definition of the customer needs in order to understand the high-level FRs and the corresponding DPs for “Care-Toy”. In order to provide the caregiver with a brief respite, the caregiver must be certain that the patient will not lose interest in the device, resulting in the patient wandering off and potentially endangering themselves. Additionally, due to the nature of the disease, patients may be extremely sensitive to unpleasant stimuli or physical restraint, limiting options for maintaining attention and preventing movement of the patient. Patients also often have impaired motor skills and cognitive abilities, meaning that any toy must be easy and comfortable to use.

The task of maintaining the attention of the patient requires both capturing their initial interest and maintaining that interest after it is captured. As a result of the symptoms of the disease, it is difficult to complete these tasks, and a different DP may be necessary to capture the interest than the DP used to retain the attention. For example, when patients lose interest in the toy and become interested in something else, an immediate external stimulus could be provided to break the patients’ diminished attention span to recapture their interest. This led to the designers creating independent modules to clarify the need for separate DPs to address the multiple FRs present.

The top-level FRs and DPs for “Care-Toy” are defined as follows. The main requirements of “Care-Toy” to provide the respite are to maintain the attention of the patient (FR1), to notify the caregiver in the case of an emergency (FR2), and to allow the patient to easily use the system (FR3). The designers conceive of a device that provides audio-visual stimulation to capture the initial interest and to hold the long-term attention of the patient (DP1). The device will have a system for detecting emergencies and notifying the caregiver (DP2), and a simple and durable physical design for ease of use (DP3). Figure 1 shows the defined high-level FRs and DPs of “Care-Toy”, as well as the relevant design matrix.

| FR       | Provides the caregiver with a brief respite (30 to 60 min) |
|----------|-----------------------------------------------------------|
| FR1      | Maintain attention of patient                             |
| FR2      | Notify caregiver in case of emergency                      |
| FR3      | Allow patient to easily use/play the toy                   |
| DP       | “Care-Toy” for the patient to interact with                |
| DP1      | Device provides audio-visual stimulation                   |
| DP2      | System for detecting an emergency and notifying            |
| DP3      | Simple and durable physical design                         |

| FR | DP1 | DP2 | DP3 |
|----|-----|-----|-----|
| FR1| X    | 0    | 0   |
| FR2| X    | X    | 0   |
| FR3| 0    | 0    | X   |

Figure 1. Top level FRs, DPs, and design matrix of “Care-Toy”.

2.2. Decomposition and Module-Junction Structure

An overview of the decomposed FRs and DPs is presented in this section. Table 1 shows the first 3 levels of decomposition representing the system structure of the “Care-Toy”. A module-junction structure based on the design matrices for Module 1 is presented in Figure 2, representing the decision order and information flow in which key modules are defined to be developed for FR1, maintaining the
attention of the patient. The module-junction structure shows the gaze detection module is the key for this system, which is described in detail in a later section.

**Table 1. Full decomposition of FRs and DPs for “Care-Toy”.

| FR0  | Provide brief report to caregiver |
|------|-----------------------------------|
| FR1  | Maintain attention                |
| FR1.1| Display images                    |
| FR1.2| Play sounds                       |
| FR1.3| Capture short-term interest       |
| FR1.4| Keep long-term attention          |
| FR1.4.1| Allow upload of media by caretaker |
| FR1.4.2| Allow selection of uploaded media  |
| FR1.5| Measure attention                 |
| FR1.5.1| Detect face of patient            |
| FR1.5.2| Detect gaze of patient            |
| FR1.5.3| Detect inputs of patient          |
| FR1.6| Recover attention if lost         |
| FR2  | Get help in case of emergency     |
| FR2.1| Detect abnormalities              |
| FR2.1.1| Collect baseline conditions       |
| FR2.1.2| Collect current conditions        |
| FR2.1.3| Compare baseline and current conditions |
| FR2.2| Alert caregiver                   |
| FR2.3| Collect and provide relevant data |
| FR3  | Allow patient to easily use toy   |
| FR3.1| Survive deeps and minor misuse    |
| FR3.2| Allow to rest in position on table |
| FR3.3| Allow to be comfortably holdable  |
| FR3.4| Allow UI to be easily accessible  |
| FR3.5| Allow freedom of movement         |

**Figure 2. Module-junction structure for Module 1 of “Care-Toy”.

To maintain the attention of the patient through a device that provides audio-visual man-machine interface, the target device must display images to the user (FR1.1) and play sounds to the user (FR1.2) that can get the patient’s attention. As patients are unlikely or unwilling to pay attention to something when they are told to, the patient must first be attracted to the device to capture short-term interest of the patient (FR1.3), which can be done through attention-grabbing flashing with interesting sounds. The device must then keep long-term (at least 30 min) interest of the patient (FR1.4) by showing familiar or interesting images and sounds to the patient. To ensure that the patient is paying attention, the system measures the attention of the patient (FR1.5) and must attempt to regain the attention of the patient if it is going to lose attention (FR1.6).

To allow the toy to use familiar images and sounds, a personalized content module needs to be created with relevant images and sounds such as family photos and popular songs that are still vivid in patients’ long-term memory (FR1.4.1). The caregiver must also be able to collect, build, and select different images and sounds to be played in different scenarios via cloud sourcing from the family (FR1.4.2).
The Attention Module, which measures the attention of the patient, will use sensors and a camera to detect if the patient is in front of the device, looking at the device, or interacting with the device. A camera is used to determine whether the user is in front of the camera through facial recognition (FR1.5.1). The camera is also used to detect the gaze of the user (FR1.5.2). Lastly, the device will determine if the user is interacting by pressing buttons or moving the device around (FR1.5.3).

In the case that the patient’s attention cannot be recovered when lost, or if there is an emergency, the system must be able to notify the caregiver (FR2). The system must be able to detect an abnormality during use (FR2.1), alert the caregiver (FR2.2), and provide the caregiver with relevant data related to the abnormality (FR2.3).

To detect abnormalities, the system will compare the current conditions with a set of baseline conditions. To do so, the system must collect the baseline conditions during a calibration period (FR2.1.1), collect current conditions from the camera and sensors during use (FR2.1.2), and determine if the current conditions are outside of what is acceptable (FR2.1.3). In order to alert the caregiver that there is an abnormality, the device can send a page or message directly to a device held by the caregiver. From the collected data, the device can output a time-stamped report.

The last main functional requirement of the device is to be easily usable by the patient (FR3), who may have impaired motor skills. The device should have a simple and durable design such that it is able to survive drops and minor misuse (FR3.1), and it must be usable in a variety of different positions, such as being able to rest on a table during use as well as able to hold in one’s hands during use (FR3.2 & FR3.3). In any of these positions, the user must be able to easily interact with the device by pushing buttons or touching the screen (FR3.4). The device must also be portable so that the user can use it in multiple different locations (FR3.5). The size, shape, material properties, and proximity of the User Interface (UI) to the holding position all affect these FRs. Figure 3 shows the coupled design matrix for FR3 (specifically for FRs 3.1, 3.3, 3.5), illustrating the need for feedback in the process of ensuring the device is strong enough to withstand abuse without being overly cumbersome to use.

|      | DP3.2 | DP3.4 | DP3.1 | DP3.3 | DP3.5 |
|------|-------|-------|-------|-------|-------|
| FR3.2 | X     | 0     | 0     | 0     | 0     |
| FR3.4 | 0     | X     | 0     | 0     | 0     |
| FR3.1 | 0     | 0     | X     | X     | X     |
| FR3.3 | 0     | 0     | X     | X     | X     |
| FR3.5 | 0     | 0     | X     | X     | X     |

Figure 3. "Care-Toy" design matrix for FR3.

2.3. Attention Measurement and Recovery Exploration
From the module-junction diagram at the 1st level FR-DP tree, the main focus in this article for the desired system becomes the task of maintaining the attention of the user. To address this, we elaborate this module on how to attract the initial attention and how to make sure that the patient is still focused (immersed) on the content of the toy. We also elaborate on the capability of the module to adapt to the severity of the affliction and the personalized behaviors of the specific patient.

2.3.1. Measure Attention. Multiple studies have been developed on the attention and engagement measurement of attention, especially with the current progress of computational power and online learning. As for unobtrusive methods, some of the technologies that have been explored include computer vision for facial expression interpretation [6], [7], pose and user gaze tracking and speech recognition [8], [9], and even remote pulse measurement [10]. These methods can be used to detect, measure and determine the user’s attention and engagement [11]–[13] while interacting with virtual systems [14], [15]. From the literature review, an approach combining facial analysis and gaze tracking was selected. Further explanation of the module development as the core of this project is presented in Section 3.
2.3.2. Recover Attention. The basis of the development works under the assumption that the attention span of the user is quite irregular and thus can easily lead to changes and loss of focus due to either external distractions or the evolution of the affliction, as previously stated. Under this assumption, it is essential to handle a wide range of elements that stimulate the patient, thus inviting them to continue the interaction, while continually feeding back the changes on the state to undertake the appropriate steps. Achieving the task requires to take into consideration and combine two parts, which can be divided into passive and active components. The passive components are those that provide a continuity and structure that allow the patient to maintain a level of familiarity and control.

2.3.3. Content. Guaranteeing the engagement of a patient requires that the system is capable of presenting instantly attractive material or stimuli to the user. The literature exploration led to the finding that the recommended content intends to stimulate the memory of the user, while also aiding in the use and honing of their capabilities. Among the studies, we can find recommendations of Sensory stimulation [16], [17] and Multi-Sensory Stimulation (MSS) [18]. The content selection would later affect the characteristics of the required interface and the ease of use. Given that the patients’ abilities keep deteriorating as a result of the affliction, an interface with high level of complexity and abstraction is undesirable. Thus the initial content selection was limited to visual stimuli from the patient’s life for memory stimulation, and low-level, task oriented games.

2.3.4. Accessibility. Previously it was mentioned that the patient’s circumstances and situation may vary, depending not only on the initial condition of the user, but also the evolution of the affliction. Thus, it is important that the system is capable of adapting to the specific scenario, however, as previous studies have found, it is very important for the elder population that the feeling of control over the system is maintained [19]. To allow the adaptation of the system to the user, the proposed solution is to allow the manual input of material by the caregiver, building the library of memories to be presented, and allow interface personalization.

3. Attention Module

Instant generation and maintaining the attention of an Alzheimer’s patient for a period of 30-60 min is the key to the success of “Care-Toy”. The selected approach is the use of computer vision for gaze tracking and user face detection. The monitoring capabilities are at the core of the “Care-Toy” system to drive audio-visual human-machine interface to get instant initial attraction, provide various content to maintain attention and notify the caregiver in case that the user’s safety may be compromised.

3.1. Module Logic.

In this subsection, we present the logic and corresponding flow diagram of the Attention Module. The description of the components for the working of the module is presented in the subsection afterwards.

The Attention Module takes care of the monitoring of the interaction between the user and the system. To ensure the continuity of such interaction, the system must complete two main tasks: the sensing of the user and the exposure to the corresponding stimuli. The stimuli are intended to be modified depending on the user and by the active input of the caregiver or recommendation of the physician. Thus, the remaining task is the monitoring or attention tracking, which falls completely on the system. In Figure 4, we can appreciate the working of such module. The first task of the module is ensuring that the user is present for the activity to start. To do this the user should be in front of the camera integrated in the system, otherwise the use of displayed and sounded alerts are used to grasp their curiosity. Once the interaction is secured, the system allows the implementation of the activities, which may be either passive or active. One addition that the activities received after early designs was the implementation of a calibration activity for the gaze tracking.
During the activities, the gaze of the user is monitored to determine if it falls within the predetermined area of interest. The monitoring of the instances when the gaze is lost may allow for the feedback to the caregiver to eliminate or substitute activities, and it may be used to monitor the progress of the patient. If the patient’s gaze is lost, the next confirming instance is the face detection. If the face is also out of range or can’t be detected, the next confirmation element is the activity sensing, which is comprised of two parts, the button pressing and the movement sensing. These two components rely on the integrated accelerometers and touch screens or keyboard of the system.

If the attention of the patient is unable to be found after a predetermined time, the system will attempt to communicate with the patient through alerts and messages or calls. If a certain number of calls remain unanswered, the system will determine a state of abandonment, thus will attempt to contact the caregiver through the corresponding strategy. If the patient resumes the interaction, the system will resume the normal monitoring procedures.

3.2. Module Components.
The prototype uses the dlib and OpenCV Python libraries for the face detection and tracking of the face position. The facial tracking is combined with the gaze tracking. To test the initial logic implementation, the gaze tracking in early stages was performed by a third party software named GazePointer [20]. The intention for this interaction is to guarantee that the user is not only facing the system, but is also looking at the desired area. The interaction of both systems allows to implement different kinds of alerts or responses depending if the user gaze is lost, but the user is still in the vision field of the system, or if the subject has abandoned it completely.

The logic testing with the use of the GazePointer tool and the facial tracking provided an acceptable response and ease to achieve the desired task, thus proving that it was possible to work with a monocular eye tracking instead of needing to implement an additional hardware piece for the viability of the system. This led to the task of developing a gaze tracking component independent of a third party. The proposed approach is the following.

Even though the final position of the gaze is delivered by a learning model, the possible approaches of which kind of data and model to use are multiple. After some analysis of the options, the designers opted to approach the problem by building on previous modules to gather the data for the training of the model. First, the eye positions were extracted by implementing the dlib libraries and defining as an area of interest the rectangle containing the points that delimit each eye. It was decided to extract each eye independently instead of using both as a single piece, as it would not only allow to minimize the additional burdening of the system in case of face pitching, but it also allowed to test the idea of developing a system capable of tracking an estimated gaze even if the obstruction of one of the eyes
may occur. Once the area containing the eye is defined, the contour of the eye is delimited using border detection to locate the pupil, and the content in such delimited area is processed by analyzing the Gaussian concentration of pigment. The response is used to generate a black and white blob whose borders are then brushed to eliminate the area of transition. With the shape acquired, the center of the area is calculated by finding the mass center of the shape. The center is then returned as the coordinates to be used in the tracking model, being referenced with the relative and global coordinates of the shape.

The training database for this model is still being built, however, the processing of the eyes has been showing promising results when evaluating the use with various light conditions and angles, although there haven’t been enough tests made with the use of eyeglasses.

4. Summary and Future Work
The current design of assistance system is the result of the consideration of the needs of the patients and caregivers, while also focusing on the available technology. By implementing the Axiomatic Design methodology, the correct identification of the functional requirements allowed for an optimized detection and prevention of potential couplings. This approach allowed the team to identify the most relevant tools needed to complete the task without sacrificing efficiency.

The currently achieved prototype is intended to be tested by potential users, which include caregivers, patients and physicians, to evaluate additional adjustments that may be needed to maximize the quality of the User Experience (UX). In addition to the interface changes that may result from such evaluation, the team is also working on optimizing and further improving detection tools. Another instance that will require further work is the expansion of the menu for available activities for the patients.

Acknowledgements
This research has been partially funded by the Tec de Monterrey and MIT Nanotechnology Program.

References
[1] P. D. United Nations, Department of Economic and Social Affairs, “World Population Ageing 2017 - Highlights (ST/ESA/SER.A/397),” 2017.
[2] World Health Organization, “Global action plan on the public health response to dementia 2017 - 2025,” 2017.
[3] W. A. Kukull et al., “Dementia and Alzheimer Disease Incidence,” Arch. Neurol., vol. 59, no. 11, p. 1737, Nov. 2002.
[4] Alzheimer Association, “2019 Alzheimer’s disease facts and figures,” 2019.
[5] N. P. Suh, Axiomatic Design: Advances and Applications, 2001.
[6] A. Kumar and A. Agarwal, “Emotion recognition using anatomical information in facial expressions,” in 2014 9th International Conference on Industrial and Information Systems (ICIIS), 2014, pp. 1–6.
[7] S. Stöckli, M. Schulte-Mecklenbeck, S. Borer, and A. C. Samson, “Facial expression analysis with AFFDEX and FACET: A validation study,” Behav. Res. Methods, vol. 50, no. 4, pp. 1446–1460, Aug. 2018.
[8] M. Aparna and P. Wanare, “Human Emotion recognition from speech,” Int. J. Emerg. Technol. Adv. Eng., vol. 4, no. 7, pp. 74–78, 2014.
[9] R. A. Calvo and S. D’Mello, “Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications,” IEEE Trans. Affect. Comput., vol. 1, no. 1, pp. 18–37, Jan. 2010.
[10] M.-Z. Poh, D. J. McDuff, and R. W. Picard, “Non-contact, automated cardiac pulse measurements using video imaging and blind source separation,” Opt. Express, vol. 18, no. 10, p. 10762, May 2010.
[11] H. Monkaresi, N. Bosch, R. A. Calvo, and S. K. D’Mello, “Automated Detection of Engagement Using Video-Based Estimation of Facial Expressions and Heart Rate,” IEEE Trans. Affect. Comput., vol. 8, no. 1, pp. 15–28, Jan. 2017.
[12] S. Dharmija, “Learning based visual engagement and self-efficacy,” in 2017 Seventh International
Conference on Affective Computing and Intelligent Interaction (ACII), 2017, vol. 61, pp. 581–585.

[13] B. M. Booth, A. M. Ali, S. S. Narayanan, I. Bennett, and A. A. Farag, “Toward active and unobtrusive engagement assessment of distance learners,” in 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII), 2017, vol. 2018-Janua, pp. 470–476.

[14] J. Hernandez, Z. Liu, G. Hulten, D. Debarr, K. Krum, and Z. Zhang, “Measuring the engagement level of TV viewers,” 2013 10th IEEE Int. Conf. Work. Autom. Face Gesture Recognition, FG 2013, 2013.

[15] J. Whitehill, Z. Serpell, Y. C. Lin, A. Foster, and J. R. Movellan, “The faces of engagement: Automatic recognition of student engagement from facial expressions,” IEEE Trans. Affect. Comput., vol. 5, no. 1, pp. 86–98, 2014.

[16] B. S. Strøm, S. Ytrehus, and E. K. Grov, “Sensory stimulation for persons with dementia: A review of the literature,” J. Clin. Nurs., vol. 25, no. 13–14, pp. 1805–1834, 2016.

[17] J. M. Witucki and R. S. Twibell, “The effect of sensory stimulation activities on the psychological well being of patients with advanced alzheimer’s disease,” Am. J. Alzheimers. Dis. Other Demen., vol. 12, no. 1, pp. 10–15, 1997.

[18] L. A. Wareing et al., “A randomized controlled trial of the effects of multi-sensory stimulation (MSS) for people with dementia,” Br. J. Clin. Psychol., vol. 40, no. 1, pp. 81–96, 2001.

[19] M. Zallio, D. Berry, and L. J. Leifer, “Meaningful Age-Friendly Design. Case Studies on Enabling Assistive Technology,” 2020, pp. 779–790.

[20] S. Deja, “GazePointer.” GazeRecorder, 2019.