Design vocabulary for human-IoT systems communication

Citation for published version (APA):
Chuang, Y., Chen, L., & Liu, Y. (2018). Design vocabulary for human-IoT systems communication. In *CHI 2018 - Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems: Engage with CHI* (Vol. 2018-April). Association for Computing Machinery, Inc. https://doi.org/10.1145/3173574.3173848

Document license:
CC BY

DOI:
10.1145/3173574.3173848

Document status and date:
Published: 20/04/2018

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 18. May. 2024

Design Vocabulary for Human–IoT Systems Communication

Yaliang Chuang
Eindhoven University of Technology, Eindhoven, The Netherlands,
yaliang@ijdesign.org

Lin-Lin Chen
Eindhoven University of Technology / National Taiwan University of Science and Technology, Taipei, Taiwan,
l.chen@tue.nl

Yoga Liu
NTU IoX Center,
Taipei, Taiwan,
liuyoga.work@gmail.com

ABSTRACT

Digital devices and intelligent systems are becoming popular and ubiquitous all around us. However, they seldom provide sufficient feed-forwards and feedbacks to reassure users as to their current status and indicate what actions they are about to perform. In this study, we selected and analyzed nine concept videos on future IoT products/systems. Through systematic analysis of the interactions and communications of users with the machines and systems demonstrated in the films, we extracted 38 design vocabulary items and clustered them into 12 groups: *Active*, *Request*, *Trigger functions*, *Approve*, *Reject*, *Notify*, *Recommend*, *Guide, Show problems*, *Express emotions*, *Exchange info*, and *Socialize*. This framework can not only inspire designers to create self-explanatory intelligence, but also support developers to provide a language structure at different levels of the periphery of human attention. Through the enhancement of situated awareness, human–IoT system interaction can become more seamless and graceful.

Author Keywords

Intelligibility; Vocabulary; Feedback; Understanding; Internet of Things.

ACM Classification Keywords

H.5.2 User Interfaces: User-centered design.

INTRODUCTION

The last few years have seen the maturation and popularization of connected products and services. With the wide spread of ubiquitous digital devices and intelligent systems, we are surrounded by many invisible agents that are continuously sensing our activities and movements [30]. Functioning together with machine learning and automated technology, the systems can also anticipate users' needs and proactively assist with their daily activities in advance [e.g., 19]. In addition, many tools have been developed for users to create services themselves [e.g., 6, 51]. These toolkits provide extensible and low-cost components for users to customize their smart homes continuously to fit their needs.

When the systems run well, they can provide users with effortless experiences, as if by magic. However, when something unexpected happens, users often feel frustrated because they don't know how the problem occurred [25, 47]. Although wireless technology provides flexibility for deploying components and extending services, unfortunately the lack of physical affordances often leads to usability issues [12].

To facilitate communication with the user, most of the designs are equipped with small embedded light emitting diodes (LEDs) to provide information and feedback; however, the lighting behaviors are often either ambiguous or unintuitive for understanding [15]. As a result, users are unable to obtain useful feedbacks to create an appropriate conceptual model. Users often don't know what the machine has sensed and what actions they should take. Thus, due to the lack of communication, many simple operations may fail [32].

In recent years, several researchers have investigated the design of good feedbacks [e.g., 15, 33] for user–product interaction. However, not many studies have addressed the challenges in communicating with a system that comprises multiple artefacts, such as a smart home. In the study presented in this manuscript, we look at the problem at a higher level to focus on the informational vocabulary that is essential for user–system communication. We started our investigation with an analysis of high-quality concept videos produced by commercial companies, research institutes, independent designers, and students. We then chose nine films from the 42 collected to analyze the human–machine, human–system, and machine–machine communications

CHI 2018, April 21–26, 2018, Montreal, QC, Canada

ACM ISBN 978-1-4503-5620-6/18/04···$15.00
https://doi.org/10.1145/3173574.3173848

depicted. The contribution of this study is the proposal of a vocabulary of 38 design terms clustered in 12 categories. We also present design implications for how this vocabulary can be applied to create situated awareness, convey information at different levels of the periphery of human attention, and design seamless interactions.

LITERATURE REVIEW

The concept of the Internet of Things (IoT) can be traced back more than three decades to Mark Weiser's [50] vision of ubiquitous computing. Fourteen years later, science fiction author Bruce Sterling [35] used the term and envisioned its usage in his futurist work *Shaping Things*. In general, the IoT will "allow people and things to be connected Anytime, Anyplace, with Anything and Anyone, ideally using Any path/network and Any service" [37: 41]. One of the early examples is the use of RFID tags that are embedded in everyday objects to provide connections between the physical and virtual worlds [49]. A user could tap the things on his/her mobile device to access the information. With the development of technology, the IoT comprises more and more components (e.g., sensors, processors, actuators) and capabilities (e.g., data collection, analysis, sharing, exchanging) to anticipate human users' requirements and react in the best manner without explicit instructions. These features formed the "smart objects" used in [5]. Compared with traditional computers, such digital artefacts can generate information about themselves and the context within a local environment. This supports the development of context-aware applications and facilitates the implicit interactions and communications with the users. Although most works on the IoT have focused on technical aspects, recently there have also been a few studies addressing the human-interaction aspects of smart products and systems. Jonell and Lopes [20] added animating behaviors to IoT artefacts as affordances to indicate their status and how the user should interact with them. For example, a door handle will change its shape to prompt a user to knock at the door before opening it and interrupting a meeting. In addition, some scholars also add the aspects of autonomy and sociability to promote the concept of the social IoT [e.g., 3]. One example is the Brad the Toaster [23] which can communicate and exchange information with other toasters on the Internet. "He" might complain that his owner does not like him anymore and ask the other objects to help him fix the relationship. Mitew [22: 13] mentioned that the sociable object is "not simply a recording device for an expanding human subjectivity, but an active participant, a mediator co-constructing the newly defined social environment."

In contrast to the abovementioned conceptual studies, Yang and Newman [47] interviewed early adaptors of the Nest Learning Thermostat to understand how this intelligent system was used in everyday life. They found that although the participants had sufficient knowledge on the advanced technology embedded in the device, due to the lack of communication, most of them still couldn't understand why the system made wrong predictions many times. This challenge was also highlighted in several studies related to the smart home where multiple artefacts were ubiquitously deployed in the environment [e.g., 12, 51]. For instance, in [51], many participants complained it was difficult to understand the system structure in terms of how the different devices were connected and coordinated to execute specific rules created by the users themselves. Those rules were also easily forgotten after being set up. Although the wireless technology helped to deploy the sensors at suitable positions, the invisibility of the signals made it difficult to check and debug the settings in triggering other devices. Without feedbacks, it was not possible to create the appropriate conceptual model [32: 138]. As a result, people could not understand why the system did not behave as they expected and they questioned its reliability and smartness.

To create mutual intelligibility and shared understanding, Suchman [36] considered that artefacts should be able to explain themselves and establish their own intelligence or rational accountability. However, when dozens of things are connected together to collaborate as an IoT system, how should the communication be designed to convey messages without overwhelming the human users? How can users be made aware of the affordances of the system and build a conceptual model? Abowd and Mynatt [1] suggested that designers could present the information at different levels of the periphery of human attention. This strategy corresponds to three types of human–computer interaction that require different levels of involvement with the IoT system, from focused interaction, to peripheral and implicit interaction [4].

In the study presented in this manuscript, we focus on the communications that are essential between a human user and an IoT product/system. Regarding intelligent properties, Suchman [36: 10] thought that smart products' behaviors are "more a matter of specifying operations and assessing their effects through the use of a common language." This implied that interaction designers should not focus only on the creation of input interfaces. Rather, they should be able to describe "what goes on between people and machines, to employ terms borrowed from the description of human interaction—dialogue, conversation, and so forth—terms that carry a largely unarticulated collection of intuitions about properties common to human communication and the use of computer-based machines." Through adding this expressive ability, the artefacts not only could express themselves in an easy understandable way, but generate graceful interaction experiences [18].

Based on these theoretical studies, we used content analysis to analyze the possible communications between a human user and an IoT product/system with videos of contemporary products or futuristic concepts. By transcribing the interactions into the English language, we collected a variety of vocabulary items that can describe the information that an intelligent product or system tries to communicate with the user or the other machines. We expect that these results can inspire developers to design feedbacks and feed-forwards that create a seamless user experience.

RESEARCH METHODS

To investigate human–system communications that will be essential in the context of future technology, an analytical approach to the concept videos was applied in this study.

First, video analysis is a popular method used by researchers and designers to understand users' perceptions and behaviors with interactive artifacts or environments in situ [e.g., 40]. Second, in recent years, many scholars have experimented with alternative video prototyping approaches to explore the functionality and contexts of new technologies or applications [e.g., 8, 43, 45]. When producing the films, the designers are able to make ideal demos that exceed natural limitations. Video prototyping offers opportunities to explore technical issues, as well as user experience or social issues. Commercial companies also produce concept videos to announce their future visions or the next generation of products. For example, in 2013, Microsoft released the "Future vision 2020" [26] concept video to demonstrate their speculative future. In the film, the designers not only designed user interactions and future gadgets, they also added many annotations to present non-visible information to convey the user experience in a way that could be easily understood by the audience. Some of the annotations might not be actually realized in the real world; nevertheless, the virtual elements in the film helped create an understanding of the user experience.

In this study, we focused on the analysis of vocabulary relating to how the machines or system would show or inform the users or other machines. We paid special attention to analyzing the annotations that designers added in the concept videos. To collect a broad vocabulary for our exploration, we used three approaches to search for related videos on the Internet. First, we used several keywords (i.e., *smart system*, *smart things*, *intelligent system*, *Internet of Things*, *IoT*, *interaction*, *user IoT communication*, and *smart home*) to search with Google and on the Youtube, Vimeo and Kickstarter websites. Second, for each result found, we also browsed the recommended videos or projects referred to by the web systems. Third, we visited the Copenhagen Institute of Interaction Design's website to investigate the students' projects. During this process, we also found some sci-fi movies related to the topics of this research; however, the clips tended to present shiny and flashy or "loud" interfaces [34]. Due to their requirement for cinematic effects and visual appeal, we did not see scenes properly portraying the continuum of interactions of a human–IoT system. Therefore, the sci-fi movies were excluded in this study.

In total we selected 42 films in which the communicative designs were explicitly demonstrated on the devices or through additional annotations to envision future scenarios. These videos covered various fields, including commercial products (14), companies' visionary videos (14), future scenarios produced by academic research institutes (5), and students' projects (9).

In order to effectively investigate various kinds of possible communication between the human user and an IoT system, we employed two approaches to analyze the 42 videos. First, one of the authors conducted an initial analysis to transcribe the explicit communications observed in each video. For instance, in "A day made of glass 2," produced by Corning Inc. [9], it was noted that because the purpose of the film was to envision ubiquitous displays in everyday life, communication was mainly focused on the presentation of different information, software functions, and telecommunications. Due to a lack of diversity in the messages conveyed within the human–IoT system, this video was not selected for further analysis. Two other visionary videos produced by Microsoft [26, 27] were excluded based on the same consideration.

Based on the results of the initial analysis, we selected nine videos (see Appendix) in which the system shown in the film conveyed more than ten different kinds of message to the human or the other machines. Then, in a second round of analysis, two other researchers were invited to analyze those nine videos further, including the implicit communications. We followed a similar approach to that used in [41] to extract information and conduct systematic analysis.

Step 1. Extraction: For each video, the three researchers independently watched and transcribed every scene that was related to human–machine/system communications. The visual content was extracted into complete sentences that included subject, verb, object, and additional contextual information. For example, the first scene in Film #8 (0:32–0:40) was transcribed as "The toy senses the child's requests and then sends commands to other machines." In this way, we could clarify the vocabulary items that would be used for showing machine-to-user or machine-to-machine communication.

Step 2. Definition: After the three researchers completed the extraction, they exchanged and discussed their data with each other. They also checked the meaning of the verbs with online dictionaries—such as the Cambridge Dictionary [7] and Vocabularies.com [46]—to incorporate their definitions.

Step 3. Classification: We then looked at the familiarity or semantic relationships among the extracted verbs. If such a relationship existed, we referred to the original scene of the concept video that the verbs were extracted from. In this way, we could achieve a better view of their connections and cluster the vocabulary items that shared similar meanings.

FINDINGS

In this study, the three researchers first independently extracted and coded 336 sentences from the nine concept videos. In the second step, the raw data were integrated as 196 sentences, with 38 unique verbs being identified. Finally, the vocabulary items were classified into 12 groups. The

No.	***Cluster***	***Vocabulary items included***	***To user***	***To machine***	***Sum***
1	**Active**	Booting, Broadcast, Processing, Join, Looking for, Monitoring, Self-check	2	19	23
2	**Request**	Request	2	3	5
3	**Trigger functions**	Activate specific modes, Send commands	4	33	37
4	**Approve**	Confirm, Identify, Recognize, Respond, Sense, Show response	47	14	61
5	**Reject**	Reject	0	1	1
6	**Notify**	Notify, Reminder, Report, Show progress	24	8	32
7	**Recommend**	Recommend, Suggest	5	0	5
8	**Guide**	Navigate, Preview	3	3	6
9	**Show problems**	Show low-energy state	1	0	1
10	**Express emotions**	Express emotion	4	2	6
11	**Exchange info**	Exchange, Negotiate, Receive, Synchronize, Transmit data, Update data	1	13	14
12	**Socialize**	Greeting, Hello, Thanks, Welcome, I see you seeing me	3	2	5

Table 1. The results of the classification of vocabulary items extracted from the videos. The numerical data represent the cumulative observations of the included terms in the extracted sentences. The *to user* and *to machine* columns indicate the main subjects with which the communications took place in the specific scenes of the videos.

resulting clusters and terms are shown in Table 1. Some examples of the specific communications are collected in Table 2. In the following sections, we will describe each category and the vocabulary items included within it.

1. Active

Definition: *busy with a particular activity; involved in a particular activity* [7]. There are seven vocabulary items clustered in this group: *Booting*, *Broadcast*, *Processing*, *Join*, *Looking for*, *Monitoring,* and *Self-check*. The machines express a status indicating that they are performing tasks. For example, in Film #1 (0:15–0:24), the animated annotation visualizes that the machine is booting. This gives the user a clear indication of the system's status and facilitates preparation to take appropriate actions afterward, such as in the film where the actor triggers the entertainment mode when he realizes the system is ready for another operation.

2. Request

Definition: *the act of politely or officially asking for something* [7]. In this study, we observed three different kinds of interaction: users request help from machines (Film #9, 2:30–2:35), machines request help from users (Film #9, 2:02–2:10), machines request help from other machines (Film #2, 0:04–0:12). In these cases, it is not enough to provide sufficient information to assist the user's activity; the designers also need to create a polite, emotional quality to the interactions.

3. Trigger functions

Definition: *to cause something to start* [7]. In Films #5 and #7, in which the producers promote an automation concept, we can see many examples of some specific mode or function being triggered by the user or events. For example, in Film #5 (2:23–2:37), when the system detects that Dad has arrived at the office, it activates the *Dad at Work* mode to automatically set up the working environment and check upcoming events. What if someday Dad enters his office and finds that the shade or computer is not turned on? In this case, Dad would need to check his App to debug the errors. By contrast, in Film #7 (0:33–0:36), the annotated beams help the user to easily figure out the problem when something unexpected happens. Although the artificial beams might not be able to be widely realized in the physical world in the near future, the concept is useful if it helps designers realize that providing clues to help users build appropriate conceptual models will improve the quality of the interaction design.

4. Approve

Definition: *to accept, allow, or officially agree to something* [7]. The main meaning of this group is providing feedbacks to the user or other machines and informing them that the expected results were accepted. For example, in Film #5 (4:35–4:39), the system identifies the grandparents and approves their entering the house. There are six vocabulary items included in this group: *Confirm, Identify, Recogonize, Respond, Sense*, and *Show response*. In such situations, it is necessary to provide an immediate response to approve or redirect the user's request. In the latter case, the designer also needs to provide appropriate explanations or suggestions to help them solve the problem.

5. Reject

Definition: *to refuse to accept, use, or believe something or someone* [7]. In this study, only one example of this category was observed, in Film #4 (1:33–1:44). Like the "Show problems" group, these somehow negative scenarios were usually eliminated in the perfect demo of concept videos [42]. In real life, bad experiences encountered by users are rejected by systems every day. However, they seldom provide reasons behind the decisions. If designers could design feedback on the lines of "Sorry, this is a problem beyond my reach," the user should feel much better.

6. Notify

Definition: *to officially give someone a piece of information* [46]. Based on the definition, the following four vocabulary items were clustered in this group: *Notify*, *Reminder*, *Report*, and *Show progress.* In addition to conventional notifications sent to the user's mobile device (such as in Film #5), the ideas in Film #2 (0:17–0:20) demonstrate the possibilities for creating more natural interactions with expressions that could be directly perceived on the devices.

7. Recommend

Definition: *to express a good opinion of; to make attractive or acceptable* [46]. Recommendation is one of the main applications for intelligent products or systems. As the things surrounding us become "smart," users might feel overwhelmed by scenarios such as the enormous advertisements shown in web browsers. In Films #4 (0:25–0:31) and #6 (0:22–0:24), the designers imitate human social behaviors to show suggestions and discuss the options with the user. The key is to enable users to have the freedom to make their own decisions.

8. Guide

Definition: *to show someone how to do something difficult or how to get somewhere* [7]. One of the advantages of IoT is the integration of cyber systems and the physical world. The location based service is one example of such applications, such as GPS navigation. For instance, in Film #7 (0:24–0:27), the designers envision virtual projected lights to guide the user in finding a specific product in the store. In Film #9 (2:37–2:52), the system not only navigates for the child to find his cat, but reports her daily activity to guide him on how to heal her. Moreover, in In Film #8 (2:49–3:00), when the system suggests Dad to give her daughter some more time of sleeping, it checks the traffic condition and Dad's calendar and previews the feasible time for commuting. This information helps the user to understand the system's recommendations and choose the best way to fulfill both his and his daughter's needs.

9. Show problems

Definition: *to display problems of itself or the system*. For example, in Film #3 (1:18–1:22), the designer uses sleepy eyes to represent that the device has encountered the problem of low energy. This metaphor can easily be interpreted to inform the user to recharge the device. Although this is the sole case observed in all of the nine selected films, like the *Reject* term, we thought this is due to the ideal demonstration of the presentation in the videos. We can imagine that various kinds of hardware or network errors might be encountered in the real world. Therefore, how to facilitate identification of problems with a design that helps the user to solve them is an important challenge.

10. Express emotions

Definition: *to display its emotional status or reactions*. We found six examples in this category in the nine concept videos. In addition to the four cases generated by the robot (Film #6) and the virtual housekeeper (Film #4), it is interesting to see in the other two cases the idea that an ordinary machine expresses its emotions to the user or to other machines. First, in Film #2 (0:34–37), the red pen expresses disappointment because the user didn't pick it. Second, in Film #4 (1:44-—1:47), the sofa tells everybody she feels sad because the vacuum cleaner declined to clean her. Although these two examples sound like stories in a soap opera, emotion is a powerful communicative device [32]. If only designers can understand how to build emotions into machines, it will bring similar benefits to the artefacts as our emotional system provides to us, such as rapid responses to avoid danger and accident.

11. Exchanging info

Definition: *to give something to someone and receive something from that person* [7]. As the number of observations indicates, this group of vocabulary items was mainly used in scenarios involving interactions between machines. An example can be found in Film #2 (0:50–0:55) where the pencil exchanges information with the book. There are six vocabulary items grouped in this category: *Exchange*, *Negotiate*, *Receive*, *Synchronize*, *Transmit data*, and *Update data*. As more and more everyday objects become smart devices, in the future new interactive behaviors among them might be developed, such as *Negotiate* as observed in Film #9 (2:00–2:04).

12. Socialize

Definition: *to take part in social activities; to interact with others* [46]. The addition of a social aspect to the IoT is also a new trend in the research community [3]. From the concept videos, we observed five cases in which machines apply social vocabulary to communicate with people or other machines. There are five vocabulary items in this cluster: *Greeting*, *Hello*, *Thanks*, *Welcome*, and *I see you seeing me.* The term *Thanks* was observed in Film #2 (0:11–0:13). The plants express their appreciation to the lamp. The other interesting term *I see you seeing me* was mainly found in the human–robot interaction in Film #6 (e.g., 2:18–2:22). These designs help build perceptual crossing and engage users' involvement [11]. By utilizing this social vocabulary, designers can improve the design qualities of user–system interactions.

	Category	***Vocabulary items included***	***Video No.***	***Time***	***Example Video Links***
1	Active	Booting	1	0:15–0:24	https://youtu.be/FnQb0y3Eijo?t=14s
		Broadcast	1	0:35–0:38	https://youtu.be/FnQb0y3Eijo?t=35s
		Processing	9	1:09–1:36	https://youtu.be/RDgTjYb2MBI?t=1m9s
		Join	8	0:59–1:04	https://youtu.be/I7VrsyqMnkU?t=59s
		Looking for	6	1:11–1:14	https://youtu.be/3N1Q8oFpX1Y?t=1m10s
		Monitoring	6	0:19–0:22	https://youtu.be/3N1Q8oFpX1Y?t=19s
		Self-check	4	1:15–1:20	https://youtu.be/i5AuzQXBsG4?t=1m15s
2	**Request**	Request	4	1:33–1:37	https://youtu.be/i5AuzQXBsG4?t=1m32s
3	**Trigger functions**	Send commands	8	0:34–0:39	https://youtu.be/I7VrsyqMnkU?t=33s
		Activate specific modes	7	1:06–1:10	https://youtu.be/-1Cgax65ctY?t=1m6s
4	**Approve**	Confirm	9	2:10–2:13	https://youtu.be/RDgTjYb2MBI?t=2m7s
		Identify	5	4:35–4:39	https://youtu.be/NjYTzvAVozo?t=4m34s
		Recognize	5	0:53–0:56	https://youtu.be/NjYTzvAVozo?t=51s
		Respond	2	0:10–0:12	https://youtu.be/QkLty3bzY4A?t=6s
		Sense	3	0:56–1:06	https://youtu.be/HZzPVAKikmo?t=54s
		Show response	9	2:30–2:35	https://youtu.be/RDgTjYb2MBI?t=2m29s
5	**Reject**	Reject	4	1:33–1:44	https://youtu.be/i5AuzQXBsG4?t=1m39s
6	**Notify**	Notify	2	0:17–0:20	https://youtu.be/QkLty3bzY4A?t=17s
		Reminder	2	0:46–0:49	https://youtu.be/QkLty3bzY4A?t=46s
		Report	4	0:50–1:03	https://youtu.be/i5AuzQXBsG4?t=50s
		Show progress	5	2:27–2:37	https://youtu.be/NjYTzvAVozo?t=2m27s
7	**Recommend**	Recommend	2	0:34–0:37	https://youtu.be/QkLty3bzY4A?t=34s
		Suggest	4	0:25–0:31	https://youtu.be/i5AuzQXBsG4?t=26s
8	**Guide**	Navigate	7	0:24–0:27	https://youtu.be/-1Cgax65ctY?t=21s
		Preview	8	2:49–3:00	https://youtu.be/I7VrsyqMnkU?t=2m49s
9	**Show problems**	Show low-energy state	3	1:18–1:22	https://youtu.be/HZzPVAKikmo?t=1m17s
10	**Express emotions**	Express emotion	6	2:22–2:27	https://youtu.be/3N1Q8oFpX1Y?t=2m25s
11	**Exchange info**	Exchange	2	0:50–0:55	https://youtu.be/QkLty3bzY4A?t=50s
		Receive	8	1:12–1:15	https://youtu.be/I7VrsyqMnkU?t=1m12s
		Synchronize	7	0:51–0:54	https://youtu.be/-1Cgax65ctY?t=51s
		Transmit data	1	0:47–0:57	https://youtu.be/FnQb0y3Eijo?t=47s
		Update data	5	4:13–4:17	https://youtu.be/NjYTzvAVozo?t=4m13s
		Negotiate	9	2:02–2:10	https://youtu.be/RDgTjYb2MBI?t=2m2s
12	**Socialize**	Hello	8	1:00–1:02	https://youtu.be/I7VrsyqMnkU?t=1m
		Thanks	2	1:11–1:13	https://youtu.be/QkLty3bzY4A?t=11s
		Welcome	4	2:54–3:02	https://youtu.be/i5AuzQXBsG4?t=2m54s
		Greeting	6	1:33–1:40	https://youtu.be/3N1Q8oFpX1Y?t=1m33s
		I see you seeing me	6	2:18–2:22	https://youtu.be/3N1Q8oFpX1Y?t=10s

Table 2. Examples of the specific communications observed in the videos.

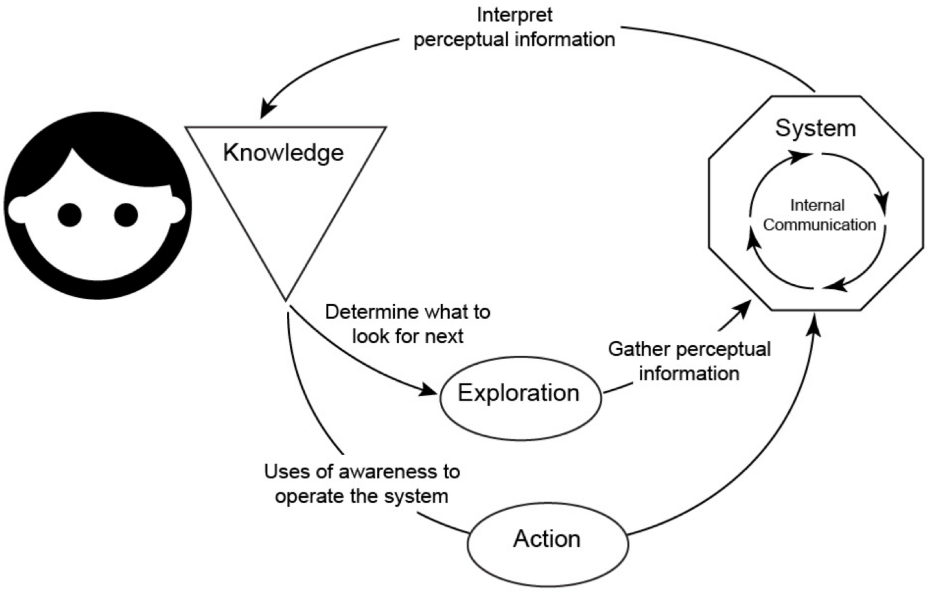

Figure 1. The awareness and interaction framework adapted from [14, 31].

DISCUSSION

Regarding communication behavior, the transmitter/receiver interaction model is widely applied in design. It was based on a telecommunications model that describes the relationship of the initiator, channel, and receiver of a message [39]. Neisser [29] added the time and spatial dimensions and proposed a perceptual cycle to represent human beings' continuous visual perception activity in gathering information from the environment, anticipating meanings and seeking more information about the situation. In this model, human perception is a constructive process for obtaining more information. Based on Neisser's concept, Gutwin and Greenberg [14] extended it to a perception–action cycle to identify human beings' awareness of the context and their actions for tackling specific events (Figure 1). Although their awareness framework was originally used to represent social collaborations in the workspace, we thought that due to the reactive, purposeful, and social properties of the IoT product/system [36], it is also suitable for describing the human–IoT system communications observed from the analysis of the concept videos. Therefore, we adapted it into an awareness and interaction framework (Figure 1) to discuss how the vocabulary terms could be used to facilitate human–IoT system communication.

In this model, a user will continuously receive and interpret perceptual information from the system operating in the environment. During this process, he/she might consciously or unconsciously look at the environment to gather more perceptual information on the system or internal machine–machine communications [31]. He/she might also initiate a specific action to operate the system. Bakker and Niemantsverdriet [4] defined three types of interactions to describe how people perform such actions with varying levels of attention. First is the *implicit interaction* through which an automatic system could detect a user's unconscious behavior (e.g., presence or movement) and trigger specific functions. For example, in Film #7 (*In Sync with life*, 0.33–0:36), the office system detects that the worker is leaving the office and changes the office to *Off* mode to save energy. In this case, we notice that it is essential to give the worker a *confirmation* clue to anticipate the upcoming changes, such as the "*ding*" auditory feedback used in the film. Without it, he would not know whether the system had sensed his action or not. If he thought the system had failed to recognize his intentions, he would switch to the *focused interaction* by waving his arm in front of the motion sensor, or open his mobile app to debug the errors. Through this example, we can see that human–IoT communication is more like a continuum constructed during the interaction flow. Between those two types of interaction, there is *peripheral interaction* performed habitually and to some extent subconsciously [3: 2]. Most designs of ambient display deliver information in a subtle manner so that it can be perceived by the user with peripheral attention. In the same video, #7 (0:53), we see the three lamps *synchronize* with a TV program and simultaneously blink to celebrate an exciting moment with the user.

On the other hand, if we compare the results with currently available products or previous research, such as the five feedbacks of Amazon Echo [2], the 17 designs of Google Home [17], or the five information states concluded in Harrison et al.'s study [15], we find that our results not only provide sufficient coverage, but also contribute possible scenarios for future designs. For instance, Mitew [22] forecasted that in the next wave of the IoT, there would be more and more social objects entering our everyday life.

Such sociable objects would not simply compute data, but would also play an active role in the newly defined social environment. In our study, we identified five social vocabulary items ("*Greeting*, *Hello*, *Thanks*, *Welcome*, *I see you seeing me*") that were absent from Amazon Echo and Google Home. As the interactive device developed in [11] demonstrated, such social feedbacks could increase users' understanding and engagement.

Design Implications

Based on the awareness and interaction framework (Figure 1), we think that the collections of design vocabulary proposed in this study can contribute to designing the interactions of human–IoT systems through the following three aspects. First, the vocabulary framework could be used as a checklist for evaluating designs. For example, the three vocabulary clusters *Active*, *Approve*, and *Trigger functions* could guide designers to check the completeness of feedbacks in enhancing a user's awareness of the system through *implicit* or *peripheral interaction*. When the system needs users' attention and focus, the designer could also create alternative feedbacks by looking at the following clusters: *Request*, *Notify*, *Recommend*, *Guide,* and *Show problems.* Such meaningful feedbacks could provide informative clues that help users build an appropriate conceptual model of the system and therefore increase their enjoyment of seamless interactions with it and/or with other users.

Second, we think that the two vocabulary clusters *Trigger functions* and *Exchange info* can guide designers to develop suitable feedbacks in constructing the internal and external communications of the system. As we saw in the concept videos #4, #5, and #7, annotations of machine-to-machine communications can help users to understand what is happening within an IoT system. Due to silent or invisible operations, people often encounter usability problems with modern technologies [32]. Norman [31] thought that feedbacks on machines' or systems' internal operations are essential for awareness, reassurance, and anticipation of taking further action. Moreover, when considering socially interacting machine units, we also need external signals that make it easier for people to know what actions are intended or being performed, such as car turn signals and brake lights that tell other drivers of our actions and intentions. The distinction between internal and external communications could differentiate the messages and help users to switch their attention to specific focuses to interpret the particular meanings.

Finally, as the IoT devices and systems become collaborative agents embodied in human activity [8], the vocabulary of the *Expressing emotions* and *Socializing* clusters could help inspire developers to design expressions of intelligent autonomy and help to establish intimate human–system relationships. This bonding feature is becoming important for the design of IoT products. For instance, in [48], the participants projected human emotions onto and formed attachments with the Morse Things, which are sets of ceramic bowls and cups that communicate solely with each other over an Internet connection. Most of the users wanted to have additional feedbacks to understand their communications or even desired to join the conversations with them. The other example is robotic products for the domestic or workspace. Forlizzi [13] used ethnographic study to investigate how the robotic vacuum cleaner was used by domestic users. She found that people tended to make social attributions to the Roomba. Some users applied male names to the robot, while others reported talking to it as it did its jobs. They treated the product as a new family member with a specific personality. This tendency in human nature implied a requirement for social communications that could elicit functional, aesthetic, and symbolic social and emotional responses. One possible approach is to imitate human behaviors in the artifact. For instance, Deckers et al. [11], developed the PeP+, which simulated eye contact behavior ("I see you seeing me") to acknowledge the user's presence. It was found that this perceptive interaction not only facilitated users' engagement, but also expressed the device's intentionality and autonomy. Those two characteristics may enhance people's perception of a device's intelligent qualities [42]. Likewise, the "lighting up" interaction was appreciated by many users of the Nest Learning Thermostat [47]. Similar to the PeP+, this simple interaction increased users' awareness of the system and expressed its anticipation of actions to change indoor temperature settings through the user interface.

LIMITATIONS

In order to get an overall picture of the possible interactions with an IoT system, we purposely selected future concept videos for analysis. Although this method has not been widely used by HCI researchers, it seems to be a very effective way to understand the user experience of future technologies, especially invisible communications between humans and intelligent products or systems. To overcome the potential limitations, we collected 42 videos at the beginning of the study. This provided us with a sufficient landscape for selecting the nine videos for conducting detail analysis.

In addition, due to our decision to transcribe the visual video into written English for content analysis, the end results of the framework might inherit the characteristics of the language in nature, such as overlapping meanings or synonyms. We envisioned that a single communication might manifest multiple vocabulary clusters. There might also be overlapping of the clusters regarding the semantic meanings in some contexts that might not be considered in our analysis. In this case, it will be recommended to browse the 38 vocabulary terms and check the video examples to find proper ideas. This will help to clarify communications and prevent ambiguity or misunderstanding. Further more, like the language, the set of vocabulary might be extended or developed alone with the time.

FUTURE WORK

Currently, we are applying the vocabulary to designing the communications of a conceptual IoT system. Extending from [15], we chose visual light communication as the main modality and designed alternative lighting behavior to convey specific vocabulary for different events, such as synchronizing with other machines or triggering the other devices to wake up. We also investigate how people will understand and interpret the communications in an interaction context. Through collecting and analyzing users' free interpretations of those expressions, we aim to validate the framework of the vocabulary. It might also contribute some more specific strategies and examples on how a developer could utilize the language structure to design effective communications for a human user and the IoT system. In the future, we will also expand the investigations to explore how the vocabularies can be used to sensitize designers and inspire them to enrich the interactivity of artefacts [e.g., 21]. Rather than simply providing functional feedbacks, as described in the "Design Implications" section, the vocabularies could empower designers' imaginations in creating intelligent autonomy and intimate social relationships.

CONCLUSION

As we surround ourselves with more and more smart objects, feedback is probably even more essential when we interact with those objects than with other people. We need to know what is happening, what the machine has detected, what its status is, and what actions it is about to perform [32]. To understand the various kinds of possible situation that the user wants to know about or the machine wants to communicate, we selected and analyzed nine concept videos on human–system interactions. From these we extracted a vocabulary of 38 design terms clustered in 12 groups. The findings do not simply cover the feedback design of available systems, but express possibilities for designing future objects with sociability. With this framework, developers would be able to build a language structure to convey the artefact's intended purpose to the user, while at the same time it could also establish the rational accountability for increasing human users' situated awareness. In this way, human users could effortlessly anticipate the status of IoT systems from implicit or peripheral communications and switch their attention to have focused interactions when needed. This human-like communication could make the interaction experience more graceful and enjoyable.

ACKNOWLEDGEMENT

This research was supported in part by the Ministry of Science and Technology of Taiwan (MOST 106-2633-E-002-001), National Taiwan University (NTU-106R104045), and Intel Corporation.

APPENDIX

A list of the concept videos selected for analysis in the study presented in this manuscript.

No.	*Title*	*Producer*	*Hyperlink*
1	Multifunctional light	SONY	Link
2	Have you ever noticed that everything chatting quietly	Yung-Hsun CHEN	Link
3	Ulo	Vivien Muller	Link
4	The Social Web of Things	Ericsson	Link
5	Life Simplified with Cloud-Based Automation	Brigham Young University	Link
6	Jibo: The World's First Social Robot for the Home	Jibo	Link
7	In Sync with Life	Samsung	Link
8	Apps4Home (Kids' Edition)	Intel-NTU	Link
9	Apps4Home—Family Edition	Intel-NTU	Link

REFERENCES

1. Gregory D. Abowd and Elizabeth D. Mynatt. 2000. Charting past, present, and future research in ubiquitous computing. *ACM Trans. Comput.-Hum. Interact. 7*, 1, 29-58. DOI: http://dx.doi.org/10.1145/344949.344988
2. Amazon. (n.d.). *Echo help: About the light ring*. Retrieved Feb. 2, 2017, from https://www.amazon.com/gp/help/customer/display.html?nodeId=201601790
3. Luigi Atzori, Antonio Iera, and Giacomo Morabito. 2014. From 'smart objects' to 'social objects': The next evolutionary step of the Internet of Things. *IEEE Communication Magazine*, *Networks*, vol. 52, no. 1, 2014, 97-105.
4. Saskia Bakker and Karin Niemantsverdriet. 2016. The interaction-attention continuum: Considering various levels of human attention in interaction design. *International Journal of Design, 10*(2), 1-14.
5. Michael Beigl, Hans-W. Gellersen, and Albrecht Schmidt. 2001. MediaCups: Experience with design and use of computer-augmented everyday objects. *Computer Networks, 35*(4), 2001, 401-409.
6. A. J. Bernheim Brush, Bongshin Lee, Ratul Mahajan, Sharad Agarwal, Stefan Saroiu, and Colin Dixon. 2011. Home automation in the wild: Challenges and opportunities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '11). ACM, New York, NY, USA, 2115-2124. DOI: http://dx.doi.org/10.1145/1978942.1979249
7. Cambridge Dictionary. n.d. http://dictionary.cambridge.org/ Last access date: 16 January 2016
8. Nazli Cila, Iskander Smit, Elisa Giaccardi, and Ben Kröse. 2017. Products as agents: Metaphors for designing the products of the IoT age. In *Proceedings of the 2017 CHI Conference on Human Factors in*

Computing Systems (CHI ’17). ACM, New York, NY, USA, 448-459. DOI: https://doi.org/10.1145/3025453.3025797

9. Corning. 2012. *A day made of glass 2*. Retrieved Sept. 15, 2017, from https://youtu.be/jZkHpNnXLB0
10. Marco de Sá, Judd Antin, David Shamma, and Elizabeth F. Churchill. 2011. Mobile augmented reality: video prototyping. In *CHI ’11 Extended Abstracts on Human Factors in Computing Systems* (CHI EA ’11). ACM, New York, NY, USA, 1897-1902. DOI: http://dx.doi.org/10.1145/1979742.1979927
11. Eva Deckers, Stephan Wensveen, Pierre Levy, and Rene Ahn. 2013. Designing for perceptual crossing: Designing and comparing three behaviors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI ’13). ACM, New York, NY, USA, 1901-1910. DOI: http://dx.doi.org/10.1145/2470654.2466251
12. W. Keith Edwards and Rebecca E. Grinter. 2001. At home with ubiquitous computing: Seven challenges. In *Proceedings of the 3rd international conference on Ubiquitous Computing* (UbiComp ’01), Gregory D. Abowd, Barry Brumitt, and Steven A. Shafer (Eds.). Springer-Verlag, London, UK, 256-272.
13. Jodi Forlizzi. 2007. How robotic products become social products: An ethnographic study of cleaning in the home. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction* (HRI ’07). ACM, New York, NY, USA, 129-136. DOI: http://dx.doi.org/10.1145/1228716.1228734
14. Carl Gutwin and Saul Greenberg. 2002. A descriptive framework of workspace awareness for real-time groupware. *Comput. Support Coop. Work, 11*(3-4), 411-446. DOI: https://doi.org/10.1023/A:1021271517844
15. Chris Harrison, John Horstman, Gary Hsieh, and Scott Hudson. 2012. Unlocking the expressivity of point lights. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI ’12). ACM, New York, NY, USA, 1683-1692. DOI: http://dx.doi.org/10.1145/2207676.2208296
16. Gartner Research. (n.d.) http://www.gartner.com/it-glossary/internet-of-things/
17. Google. (n.d.). Google Home LED lights. Retrieved Feb. 2, 2017, from https://support.google.com/googlehome/answer/7073219?hl=en
18. Philip Hayes and Reddy Raj D. 1983. Steps toward graceful interaction in spoken and written man machine communication. *International Journal of Man-Machine Studies 19*: 231-284, 1983.
19. Amy Hwang and Jesse Hoey. 2012. Smart home, the next generation: Closing the gap between users and technology. In *AAAI Fall Symposium at Arlington*, Virginia, USA. 14-21.
20. Patrik Jonell and Pedro Lopes. 2015. Affordance++: The tale of animating IoT objects. *XRDS 22*, 2, 20-21.
21. Youn-kyung Lim, Sang-Su Lee, and Da-jung Kim. (2011). Interactivity attributes for expression-oriented interaction design. *International Journal of Design, 5*(3), 113-128.
22. Teodor Mitew. 2014. Do objects dream of an internet of things? *The Fibreculture Journal, 23*, 1-25.
23. Simone Rebaudengo, n.d. Addicted products: The story of Brad the Toaster. Video. Retrieved Sept. 15, 2017, from https://vimeo.com/41363473
24. Sarah Mennicken and Elaine M. Huang. 2012. Hacking the natural habitat: an in-the-wild study of smart homes, their development, and the people who live in them. In *Pervasive Computing. Pervasive 2012. Lecture Notes in Computer Science*, vol 7319, Judy Kay, Paul Lukowicz, Hideyuki Tokuda, Patrick Olivier, and Antonio Krüger (Eds.). Springer-Verlag, Berlin/Heidelberg, Germany, 143-160. DOI: http://dx.doi.org/10.1007/978-3-642-31205-2_10
25. Sarah Mennicken, Jo Vermeulen, and Elaine M. Huang. 2014. From today’s augmented houses to tomorrow’s smart homes: new directions for home automation research. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (UbiComp ’14). ACM, New York, NY, USA, 105-115. DOI: http://dx.doi.org/10.1145/2632048.2636076
26. Microsoft. 2013. Microsoft’s Concept - Future vision 2020. Video. Retrieved Sept. 15, 2017, from https://www.youtube.com/watch?v=ozLaklIFWUI
27. Microsoft. (n.d.) Productivity future vision. Retrieved Sept. 15, 2017, from https://www.youtube.com/watch?v=w-tFdreZB94
28. MIT House_n Research Group. (n.d.) The PlaceLab. http://web.mit.edu/cron/group/house_n/placelab.html
29. Ulric Neisser. 1978. Perceiving, anticipating, and imagining. *Minnesota Stud. Philos. Sci. 9*, 89-106.
30. Anton Nijholt. 2004. Where computers disappear, virtual humans appear. *Computers & Graphics, 28*, 4 (2004), 467-476.
31. Don Norman. 1992. *Turn signals are the facial expressions of automobiles*. Basic Books.
32. Don Norman. 2007. *The design of future things*. Basic Books.
33. Alice V. Pintus. 2010. Tangible lightscapes. In *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction* (TEI ’10). ACM, New York, NY, USA, 379-380. DOI: http://dx.doi.org/10.1145/1709886.1709988

34. Michael Schmitz, Christoph Endres, and Andreas Butz. 2008. A survey of human-computer interaction design in science fiction movies. In *Proceedings of the 2nd international conference on INtelligent TEchnologies for interactive enterTAINment* (INTETAIN '08). ICST, Brussels, Belgium, Article 7, 10 pages.

35. Bruce Sterling, 2005. *Shaping things*. MIT Press.

36. Lucy A. Suchman. 1987. *Plans and situated actions: The problem of human-machine communication*. Cambridge University Press.

37. Harald Sundmaeker, Partick Guillemin, Peter Friess, and Sylvie Woelffle. 2010. *Vision and challenges for realizing the Internet of Things*. Publications Office of the European Union.

38. L. Takayama, C. Pantofaru, D. Robson, B. Soto, and M. Barry. 2012. Making technology homey: Finding sources of satisfaction and meaning in home automation. In *Proc. Ubicomp 2012*. 511-520.

39. Claude Elwood Shannon & Warren Weaver. 1949. *The mathematical theory of communication*. University of Illinois Press.

40. Deborah Tatar. 1989. Using video-based observation to shape the design of a new technology. SIGCHI Bull. 21, 2 (October 1989), 108-111. DOI: http://dx.doi.org/10.1145/70609.70628

41. Sara Tena, Paloma Díaz, David Díez, and Ignacio Aedo. 2012. Bridging the communication gap: A user task vocabulary for multidisciplinary web development team. In *Proceedings of the 13th International Conference on Interacción Persona-Ordenador* (INTERACCION '12). ACM, New York, NY, USA, Article 27, 7 pages. DOI: http://dx.doi.org/10.1145/2379636.2379662

42. Serge Rijsdijk & Erik Hultink. 2009. How today's consumers perceive tomorrow's smart products. *Journal of Product Innovation Management, 26*, 24 - 42.

43. Axel Roesler, Barbara Holder, Dan Ostrowski, Nate Landes, Stephen Minarsch, Daniya Ulgen, Erin Murphy, and Haeree Park. 2017. Video prototyping for interaction design across multiple displays in the commercial flight deck. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (DIS '17). ACM, New York, NY, USA, 271-283. DOI: https://doi.org/10.1145/3064663.3064800

44. Brian L. Thomas and Diane J. Cook. 2014. CARL: Activity-aware automation for energy efficiency. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* (UbiComp '14 Adjunct). ACM, New York, NY, USA, 939-946. DOI: http://dx.doi.org/10.1145/2638728.2641554

45. Bruce Tognazzini. 1994. The "Starfire" video prototype project: A case history. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '94), Beth Adelson, Susan Dumais, and Judith Olson (Eds.), 99-105. DOI: http://dx.doi.org/10.1145/191666.191712

46. Vocabulary Dictionary. https://www.vocabulary.com/dictionary/ Last access date: 16 January 2016

47. Rayoung Yang and Mark W. Newman. 2013. Learning from a learning thermostat: lessons for intelligent systems for the home. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing* (UbiComp '13). ACM, New York, NY, USA, 93-102. DOI: http://dx.doi.org/10.1145/2493432.2493489

48. Ron Wakkary, Doenja Oogjes, Sabrina Hauser, Henry Lin, Cheng Cao, Leo Ma, and Tijs Duel. 2017. Morse Things: A Design Inquiry into the Gap Between Things and Us. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (DIS '17). ACM, New York, NY, USA, 503-514. DOI: https://doi.org/10.1145/3064663.3064734

49. Roy Want, Kenneth P. Fishkin, Anuj Gujar, and Beverly L. Harrison. 1999. Bridging physical and virtual worlds with electronic tags. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (CHI '99). ACM, New York, NY, USA, 370-377. DOI: http://dx.doi.org/10.1145/302979.303111.

50. Mark Weiser. 1991. The computer for the 21st century. *Scientific American, 265*(3), 94-104.

51. Jong-bum Woo and Youn-kyung Lim. 2015. User experience in do-it-yourself-style smart homes. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (UbiComp '15). ACM, New York, NY, USA, 779-790. DOI: http://dx.doi.org/10.1145/2750858.2806063