Towards the Unification of Computational Thinking and EUDability: Two Cases from Healthcare

Barbara Rita Barricelli
University of Brescia
Brescia, Italy
barbara.barricelli@unibs.it

Daniela Fogli
University of Brescia
Brescia, Italy
daniela.fogli@unibs.it

Luigi Gargioni
University of Brescia
Brescia, Italy
luigi.gargioni@unibs.it

Angela Locoro
University of Brescia
Brescia, Italy
angela.locoro@unibs.it

Stefano Valtolina
University of Milan
Milano, Italy
valtolin@di.unimi.it

ABSTRACT

This paper presents a study about the mapping of the EUDability of End-User Development (EUD) tools with the Computational Thinking (CT) skills of users. This mapping provides an approach to evaluate the suitability of a EUD environment in supporting people performing their daily work while managing and exploiting EUD tools. EUDability is a construct encompassing different dimensions that need to be assessed through a careful scrutiny by human-computer interaction experts, while CT skills should mirror those dimensions from the point of view of assessing the level of ability of users in managing problems with a computational thinking attitude. Moving from the healthcare domain, we present two cases: a tool for geriatric professionals supporting them in the preparation of cognitive exercises for elderly patients; and a tool for pharmacists, which empowers them to create robot programs related to the preparation of personalized medications. These cases have been exploited to show how to unify the EUDability assessment with the CT skills assessment. In particular, the application of the EUDability evaluation method for each tool, as well as the administration of the Computational Thinking Scale to domain experts are shown. The results of the two assessments are reported and discussed, together with the limitations of the present study. The results show the goodness of fit of the proposed EUD tools in the healthcare domain.

CCS CONCEPTS

• **Human-centered computing → HCI theory, concepts and models**.

KEYWORDS

End-User Development, Computational Thinking, Evaluation, Healthcare

AVI 2024, June 03–07, 2024, Arenzano, Genoa, Italy

ACM ISBN 979-8-4007-1764-2/24/06
https://doi.org/10.1145/3656650.3656671

ACM Reference Format:
Barbara Rita Barricelli, Daniela Fogli, Luigi Gargioni, Angela Locoro, and Stefano Valtolina. 2024. Towards the Unification of Computational Thinking and EUDability: Two Cases from Healthcare. In *International Conference on Advanced Visual Interfaces 2024 (AVI 2024), June 03–07, 2024, Arenzano, Genoa, Italy*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3656650.3656671

1 INTRODUCTION

End-User Development (EUD) encompasses methods and techniques to empower end users to modify, extend, and create digital artifacts [22, 28]. In [5], the authors suggest that EUD might help fostering a sustainable digital transformation, in that it could support workers to deal with the complexities entailed by the introduction of novel technologies and intelligent automation at the workplaces [21, 23]. However, EUD systems must not only be suitable to work contexts, but they must be adequate to the Computational Thinking (CT) skills of the individuals working in those contexts.

The EUDability construct has been proposed to capture the quality dimensions of EUD systems [6]. To assess each EUDability dimension, a checklist is proposed in [5], without explaining its concrete application to a real case. A first contribution of this paper consists of a simple protocol for the inspective evaluation of EUDability using that checklist.

The Computational Thinking Scale (CTS) has been proposed by Tsai et al. [33] and validated in the programming education field as a successful tool to assess the CT skills of students. It consists of a questionnaire to assess CT skills through questions related to five dimensions: Abstraction, Decomposition, Algorithmic Thinking, Evaluation, and Generalization. In this paper, we slightly revise it to make its dimensions fit those of EUDability and allow its application to problem-solving contexts, where EUD could improve effectiveness and efficiency of daily work.

We applied the above tools (EUDability checklist and CTS questionnaire) to assess two different EUD environments in the healthcare field, the former devoted to geriatric professionals who need to create and submit memory and attention exercises to older patients, the latter for pharmacists who must program a collaborative robot in the frame of personalized medication preparation.

The paper aims to demonstrate how the combined adoption of the EUDability checklist and CTS questionnaire may be useful to preliminary evaluate the suitability of a EUD environment in a specific work context, with respect to its target users. The results of the evaluation can inform a revision of the EUD environment to improve its EUDability, or, if this is unfeasible, to plan users' training sessions to make them proficient in using the EUD environment.

The paper is organized as follows: Section 2 discusses the related work about End-User Development and Computational Thinking; Section 3 presents the cases from healthcare considered in this paper for EUDability assessment; Section 4 illustrates the evaluation technique for EUDability and the results of its application to our cases; Section 5 describes the Computational Thinking Scale and the results of its administration to 20 participants; Section 6 discusses how the unification of Computational Thinking and EUDability can be carried out; Section 7 presents the limitations of the work; Section 8 summarizes the main contribution of the work and the next research steps.

2 RELATED WORK

2.1 End-User Development and EUDability

End-User Development originated in 2003 within the European Network of Excellence on End-User Development (EUD-Net). The first definition of EUD, published in [22], describes it as "a set of methods, techniques, and tools that allow users of software systems, who are acting as non-professional software developers, at some point to create, modify, or extend a software artifact". The evolution of technology led to a new definition published in [4]: "the set of methods, techniques, tools, and socio-technical environments that allow end users to act as professionals in those ICT-related domains in which they are not professionals, by creating, modifying, extending and testing digital artifacts without requiring knowledge in traditional software engineering techniques". This new description of EUD is a more modern representation of the potential offered by recent technology advancements and provides a hint at what end users can achieve when properly enabled. EUD is particularly interesting when applied to the work context in organizations where end users are professionals, i.e., domain experts, and best know the needs, challenges, and issues of their workplace and work practice. This work-oriented perspective on EUD has been addressed many times over the years and documented in the literature (e.g., [24, 26]), considering several application domains, like healthcare [11], business process management [20], and education [25]. The numerous cases described in the EUD literature share a common message: the importance of bringing end users to a greater awareness of the potential of technology, enabling them to move from being passive consumers to being active producers of software and domain-related knowledge [15]. The process to empower end users in their workplace is studied in several fields (e.g., Computer Supported Collaborative Work (CSCW) [18], Human Work Interaction Design (HWID) [1]).

To assess the validity of EUD implementations, several methods have been applied in these years, as pointed out in literature analysis works. In 2013, Paternò [27] identified the key concepts characterizing the research published in the EUD field. Among them, of particular interest, are the necessity of balancing the application complexity and the learning effort they require when designing EUD tools and the need to design the tools allowing the users to manipulate content instead of asking them to use macros, formulas, or scripting languages. As [31] pointed out, evaluations are mostly performed in laboratories instead of reaching out for action and basic research.

What emerges from the literature is that there is a lack of specialized methods for the evaluation of EUD environments. Indeed, these environments are evaluated mostly against usability and user experience with the application of general methods and techniques used in Human-Computer Interaction (HCI) [29]. Among the most well-known questionnaires used for evaluating EUD, there are the System Usability Scale (SUS) (used, for example, in [12]), the User Experience Questionnaire (UEQ) (e.g., [2]), the Computer Usability Satisfaction Questionnaire (CUSQ) (adopted by [8]), and the NASA Task Load Index (NASA TLX) (e.g., [17]). Another approach to the evaluation of EUD tools is the one belonging to Semiotic Engineering, called the Communicability Evaluation Method (CEM), used for example in [13, 34].

A method specifically designed for EUD evaluation has been proposed in [14], which analyses EUD software by using criteria related to software quality characteristics, including usability, whose evaluation is articulated according to Nielsen's principles.

The acknowledgment of the lack of a proper EUD evaluation method paved the way for the definition of the EUDability concept [6]: "EUDability is the degree of concreteness, modularity, structuredness, reusability, and testability fostered by a EUD environment designed for specified end-user developers, with a specified goal to be pursued in a specified context."

The core dimensions that characterize EUDability and that are used for its assessment are:

- *Concreteness*: the ability of a EUD environment to present concepts and requests in a concrete way, without requiring the users to possess highly-developed abstraction skills;
- *Modularity*: the availability in a EUD environment of elements, like blocks, modules or similar objects that help end users decompose a problem and identify the pieces needed to compose a solution;
- *Structuredness*: the property possessed by a EUD environment to support the creation of a step-by-step solution and to facilitate the end users to connect input and output of different steps;
- *Reusability*: the property of a EUD environment of allowing the end users to reuse and possibly share the outcome of a EUD activity;
- *Testability*: the capability of testing the outcome of a EUD activity directly in the EUD environment in which it was created.

This paper proposes a EUD inspection method based on EUDability definition and shows its application in two healthcare cases.

2.2 Computational Thinking Skills and their Assessment

Many definitions exist for the term "Computational Thinking", though all of them stem from the education domain only. For example, one of the most recent studies merged definitions from the International Society for Technology in Education (ISTE), and the Computer Science Teachers Association (CSTA), ending up with the following definition of CT [35]: "a problem-solving process that includes formulating problems, using a computer or other tools, logically organizing, analyzing, and representing data, automating solutions through algorithmic thinking, achieving efficient and effective solutions, and generalizing and transferring to other problems". Similar properties and definitions for CT were devised in [3, 19, 36], with an extension of more or less detailed tasks to be accomplished for each of the above mental processes.

Considering Computational Thinking for EUD, the following operational definition was proposed in [6]: "the acquired capability of adopting a three-stage mental process, i.e., defining the problem, solving the problem, analyzing the solution, by knowing how to apply five basic skills, i.e., abstraction, decomposition, algorithm design, generalization, and evaluation, up to an individual level of mastery.".

The design of an integrated framework where possibly to prioritize CT dimensions was proposed in [33]. After the identification of mental processes, including the five skills (i.e., "Abstraction", "Decomposition", "Algorithmic Thinking", "Evaluation" and "Generalization"), the authors propose to see CT as an integrated ability including "social-cooperative capacities (e.g., solving problems collaboratively), creative thinking (e.g., creatively formulating solutions) and critical thinking (e.g., thinking multi-dimensionally while working on problems).".

This perspective lays in the direction of going beyond a single element such as the problem-solving one, but rather considering all the elements influencing people ability to solve problems: their scholarly background knowledge, their workplace or context of routine creation and consolidation, their experience / exposure to problem solving, their individual attitudes to creativity, and the like. However, no organic systematization of these aspects have been devised yet properly, being some of them dynamic and difficult to abstract from individual and contextual factors.

In [30], the Cattell-Horn-Carroll (CHC) model considers also "fluid reasoning", which is defined as: "the use of deliberate and controlled mental operations to solve novel problems that cannot be performed automatically". These mental operations include inductive and deductive reasoning as "the hallmark indicators of fluid reasoning." [*ibidem*].

To sum up, as also remarked by Tikva and Tambouris [32], CT definitions are developing into descriptions of capabilities and skills, also including operational definitions.

The assessment of CT is currently conceived only in the education field, and may include the design and categorization of tasks [19], as well as questionnaire items assessing key dimensions such as "computational concepts", "computational practices", and "computational perspectives", regarding both the routine and the more creative activities of generalization, problem-posing, and the like [9, 37]. In particular, the Computational Thinking Scale (CTS) has been proposed for assessing CT skills in [33]. It is composed of five sub-scales corresponding to the five CT skills mentioned above (Abstraction, Decomposition, Algorithmic Thinking, Evaluation and Generalization), and is the one adopted in this study, with some minor changes introduced in Section 5.1.

3 EUD CASES IN HEALTHCARE

In this section we present two tools in the healthcare domain, the former for geriatric professionals, the latter for pharmacists.

3.1 A EUD Environment for Geriatric Professionals

Senile dementia represents one of the health emergencies we are facing now and will be facing even more in the near future due to the aging population. Various studies highlight how currently available pharmacological methodologies cannot definitively halt degenerative processes [7]. However, geriatric doctors and nurses (from now on called *geriatric professionals*) often have to devise exercises for their patients based on their theoretical and clinical knowledge and experience without any support in the time-consuming activity or possibility of sharing their competencies with other collaborators. Another problem is related to monitoring the patient's cognitive capacity. The geriatric professionals often only detect cognitive decline during periodic meetings with older people. Nevertheless, interacting promptly at the first signs of deterioration could help doctors and healthcare assistants define a monitoring and intervention plan to maintain certain functions, thus improving the quality of life for older adults. According to these considerations, this case aims at providing a web platform capable of supporting geriatric professionals in the activities listed below in the EUD field.

3.1.1 Profiling of the elderly. The primary function of the web platform is to assist geriatric professionals in profiling older adults by administering the Symptoms of Dementia Screener (S.D.S. - [16]) and the General Practitioner Assessment of Cognition (GPCog - [10]) questionnaires. These assessment tools, designed to evaluate cognitive functions in older adults, perform an initial screening to identify a prodromal or preclinical cognitive impairment stage. This identification is based on the assignment of symptoms that can already be categorized as "dementia" according to the nosographic criteria of international medicine.

3.1.2 Definition of Active Life Indices for the Elderly. The main EUD functionality of the web platform enables geriatric professionals to monitor older people's daily activities, particularly those related to the completion of specific exercises, to track the progress of their cognitive, mnemonic, and orientation abilities. The administration of exercises and the corresponding monitoring of the older adult's problem-solving skills are carried out through a specific conversational agent (chatbot) that the older adult can use on their phone or tablet. Using the platform in Figure 1, geriatric professionals can create exercises to stimulate attention, use of language, reasoning, and orientation or reuse them from a predefined repository. An appropriate module specifies the evaluation strategies the geriatric professional can use for monitoring the exercise results. It previews how the exercise will be displayed to older adults on a tablet through a conversational agent. Moreover, a geriatric professional can define each older adult's weekly personalized intervention plan (Figure 2). This way, the conversational agent can guide older adults through the exercises and help them establish a consistent exercise routine. In scheduling, it is also possible to specify when to send reminders or notifications to remind older people to perform the exercises.

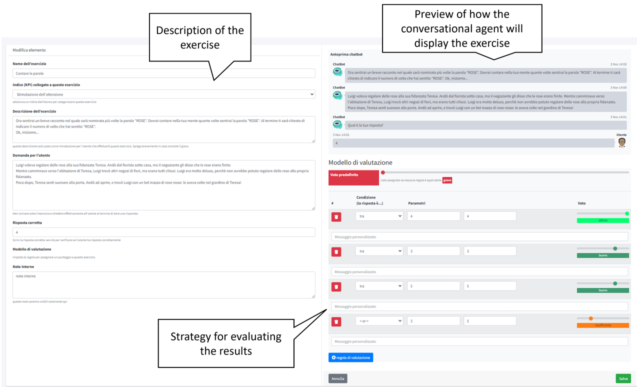

Figure 1: Creation of an exercise. On the left, the Figure represents a form for creating an exercise consisting of a title, type, description, and activity. The strategy for evaluating the results appears on the bottom right. On the top right, the dashboard previews how the conversational agent will display the exercise.

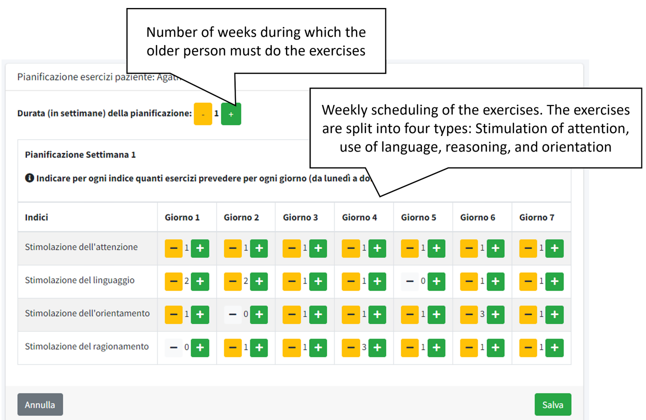

Figure 2: Weekly scheduling of the exercises. The geriatric professional specifies the number of weeks during which the older adult must do the exercises. The Figure at the bottom shows how the geriatric professional can set the weekly planning of the activities split into four types of exercises: stimulation of attention, use of language, reasoning, and orientation.

3.1.3 Configuration of the Conversational Agent. Using the platform, the geriatric professional can configure the agent that will interact with older people, specifying its gender, avatar, linguistic style, and other behavioural characteristics, such as presenting only exercises or also providing emotional support. On the site's homepage, it is possible to monitor the results of the assisted elderly individuals to assess the progress of active life indices.

3.2 A EUD Environment for Pharmacists

Galenic preparations are customized medications crafted by pharmacists or veterinarians, which serve as patient-centric remedies. These preparations enable tailored dosing, mitigating allergy concerns, cutting down expenses, and addressing the intricacies of treating rare diseases. However, the existing manual manufacturing approach for galenic preparations presents numerous hurdles for human workers.

In interviews with pharmacists, the potential benefits of using a collaborative robot in specific operational phases involved in the preparation of galenic formulations were explored. Pharmacists identified repetitive actions, tasks requiring precision, and those with low-added value as the areas most suitable for improvement. These ones pertain to the ingredient mixing phase, the capsule filling phase, and the packaging phase; all of them could be made more efficient and effective with the use of a collaborative robot.

A EUD environment based on the previous considerations was designed and developed. The environment aims to provide a natural and intuitive approach that guides pharmacists throughout robot task definition. Additionally, the user interface consistently provides comprehensive explanations of system activities. As a result of the EUD activity, a robot program is generated, which follows a structured sequence including actions such as mixing ingredients, filling capsules, and picking capsules and placing them into containers, interspersed with human actions. More precisely, the user must define the necessary details for the robot to perform each phase of the galenic preparation. These details specify the action to be taken during the mixing phase, the grid to be used during the filling phase, and the container for the packaging phase.

3.2.1 Definition of domain items. The EUD environment assists end users in defining the domain items, including objects such as grids and containers for the robot's recognition and manipulation, as well as the mixing actions to be carried out by the robot during the ingredient mixing phase. The EUD environment exploits image recognition algorithms, natural language processing, and graphical interfaces that enable pharmacists to instruct the robot regarding domain concepts. Using these interfaces, the user can define the details of each item, such as the name, action speed, number of grid slots to fill, and provide robot related data, such as container position in robot coordinates. It is also possible to define synonyms for each item to enrich the personal dictionary used by the natural language interface. Finally, when defining a new grid or container, image recognition is used to identify the element by its outline and recognise it during task execution.

3.2.2 Definition of robot tasks. After domain objects and actions have been defined, the EUD environment supports robot task definition for galenic preparations. Defining a preparation involves determining which item to use for each phase. Each preparation, which represents a robot task, is defined through a hybrid interaction paradigm. This involves a chat-based interface (Figure 3) and a graphical interface (Figure 4). In the chat-based interface, the user can indicate which previously defined items to use (in the example of Figure 3, the user is asking to use the already defined mixing action "blending"), or can define new items through the chat itself (in the example, the user is asking to use a new grid with 15 rows and 15 columns, by providing a photo of it). In the graphical interface, the user can inspect the details of the chosen items for the robot task (for instance, in Figure 4 the user can see the speed

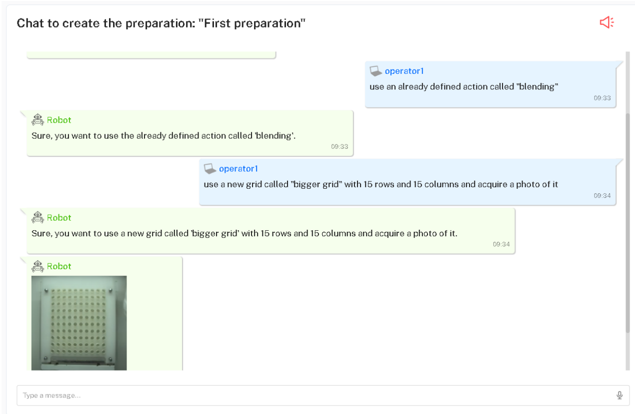

Figure 3: The chat in the EUD Environment for Pharmacists

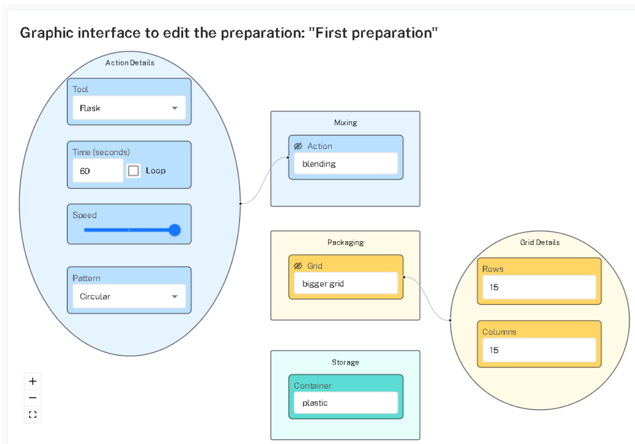

Figure 4: The graphic interface in the EUD Environment for Pharmacists

and movement pattern adopted by the "blending action"), modify item parameters, or selecting alternative items.

This hybrid approach allows the user to define a robot task intuitively using the chat, which integrates a generative Artificial Intelligence approach that leverages Large Language Models, and then validate or modify the output of the conversation using the graphical interface. Here, the robot task is represented as a sequence of blocks enabling non-expert users to assess robot program accuracy and modify blocks or parameters as needed. This provides the pharmacist with a clear and immediate representation of what has been defined.

4 EUDABILITY ASSESSMENT OF THE EUD ENVIRONMENTS

4.1 EUDability Inspection Technique

For assessing the EUDability of a EUD environment, a predictive inspection-based evaluation method can be used. The method is meant to be carried out by experts in Human-Computer Interaction and specifically in End-User Development. The evaluation consists of exploring the entire EUD environment and filling in the 15-item checklist shown in Table 1. The goal of this checklist is to measure the degree of Concreteness, Modularity, Structuredness, Reusability, and Testability fostered by a EUD environment, hence to assess the EUDability of the environment by observing its five dimensions. Once the experts complete the checklist by assigning Yes (Y) or No (N) to all items individually, they meet in a debriefing session to discuss and reach an agreement on their assignments. Then, the EUDability assessment is performed by assigning a value on the scale *{None, Low, Mid, High}* to each dimension, according to the number of Y answers provided by the HCI experts to the single items (*None* in case of no Ys, *Low* in case of a single Y, *Mid* with two Ys, and *High* with all Ys).

4.2 Applying the EUDability Checklist to Our Healthcare Cases

Three HCI experts applied the EUDability inspection technique to the applications presented in Section 3. Since the evaluators did not participate in the design of the applications, before doing the evaluation, the developers provided them with a brief description of the goals of the applications, the tasks they supported, and the target users. Then, the evaluators individually conducted their evaluation through the checklist. Finally, the evaluators met in a debriefing session where they discussed their individual evaluation and agreed on the levels to be assigned to each EUDability dimension.

4.2.1 EUDability of the environment for geriatric professionals. The results of the EUDability evaluation are the following:

- The **Concreteness** dimension has been assessed as *Mid*, since a few terms used in the application do not recall familiar elements of the experts' domain. For instance, during the creation of an exercise the geriatric professional must select a disease indicator, which is called KPI in the application, but KPI does not belong to the geriatric professional's language.
- The **Modularity** dimension has been assessed as *High*, since no issues emerged with respect to the identification, organization, and retrieval of the elements necessary to create the exercises.
- The **Structuredness** dimension has been assessed as *Mid*. In fact, even though the steps to be performed to create an exercise and to define their order are clearly suggested by the application, the connectenedness between the steps is not evident, since the application automatically generates the chatbot dialogue, without any possibility for the user to intervene.
- The **Reusability** dimension has been assessed as *Low*, since the chatbot resulting from the EUD activity automatically selects and proposes the exercises created by the geriatric professional to the patient, without giving the possibility to control the chatbot's behavior at runtime. In addition, exercises cannot be modified to be used for new patients without changing the same exercises provided to previous patients.
- The **Testability** dimension has been assessed as *Low*, since, even though a preview of the chatbot behavior is presented,

Table 1: Checklist for EUDability predictive inspection-based evaluation.

Dimensions	Items to check (Y/N forced choice)
Concreteness	The EUD environment: 1. represents the domain concepts as elements for problem solving 2. does not require to know the low-level details of each element 3. uses the language that is more familiar to the end user
Modularity	The EUD environment allows: 4. the identification of elements that may compose a problem solution 5. organizing the elements in meaningful categories 6. freely exploring each category of elements
Structuredness	The EUD environment suggests: 7. the steps to be performed for problem solving 8. the order of the steps to be performed for problem solving 9. the connectedness between different steps of the problem solution
Reusability	The outcome resulting from a EUD activity can be: 10. saved in a library 11. included in another EUD project 12. modified to be adapted to another EUD project
Testability	The outcome resulting from a EUD activity can be: 13. executed within the EUD environment 14. inspected by the user 15. represented in alternative ways

the user can no more modify it, and there are no alternative ways to see the outcome of the EUD activity, beyond the chatbot.

4.2.2 EUDability of the environment for pharmacists. The results of the EUDability evaluation are the following:

- The **Concreteness** dimension has been assessed as *High*, since the main elements refer to domain concepts, the user does not have to know the low-level details related to robot configuration and robot movements, and the terms used to name the elements belong to the pharmacists' language.
- The **Modularity** dimension has been assessed as *High*, since the elements to create robot tasks for a galenic preparation can be easily identified, retrieved and freely explored.
- The **Structuredness** dimension has been assessed as *Mid*, since the chat and graphical interfaces are organized in three steps necessary to create a robot task, and these are shown in the interface in a given order, and progressively updated; however, it is unclear if these steps are connected one another.
- The **Reusability** dimension has been assessed as *High*, since the robot tasks (called "preparations" in the application) can be saved, re-used and modified to create different robot tasks.
- The **Testability** dimension has been assessed as *Mid*, since the execution of the robot task cannot be simulated within the EUD environment, but the graphical interface can be used to inspect the robot task and possibly modify it, and it is an alternative way to the chat to represent the robot tasks.

5 THE COMPUTATIONAL THINKING SCALE AND ITS APPLICATION IN HEALTHCARE CASES

5.1 Computational Thinking Scale

To assess the CT level of end users, we decided to adopt the Computational Thinking Scale by Tsai et al. [33], which has been validated in computer programming education, but has been conceived to be as more general as possible to be rapidly adapted to any kind of problem solving. The CTS is administered to the users by means of a questionnaire including 19 items; in the Tsai's version of the CTS, the items represent statements about problem solving attitudes with computer programming, and participants are required to self-evaluate each item using a 5-point Likert scale ranging from *not agree at all* to *totally agree*.

We adapted the questionnaire by reformulating the introduction to the items, changing it from "When I solve a problem (by computer programming)..." [33] to "When I try to solve a problem in my work...". To remain consistent with the EUDability dimensions, we also swapped the order of the fourth and fifth sub-scales (Evaluation and Generalization, respectively), and named "Algorithm Design" instead of "Algorithmic Thinking" the third one, because the latter is almost synonymous with Computational Thinking. We finally decided to use a 6-point Likert scale to avoid the central tendency bias. Our version of CTS is presented in Table 2.

Tsai et al. [33] determine the results of CTS calculating the mean value for each sub-scale. Our version of CTS and its original application allows making a step further, i.e. determining the users' level of each CT skill and assigning to this level a value on the scale *{Low, Mid, High}* (see Section 5.2).

Table 2: The CTS skills and items for CT level assessment (adapted from [33]).

When I try to solve a problem in my work ...
Abstraction
AB-1: I usually think of a problem from a whole point of view, rather than looking at the details
AB-2: I usually think about the relations between different problems
AB-3: I usually try to find the key points of a problem
AB-4: I usually try to analyze the common patterns of different problems
Decomposition
DE-1: I usually think if it is possible to decompose a problem
DE-2: I usually think of the structure of a problem
DE-3: I usually think about how to split a big problem into several small ones
Algorithm Design
AD-1: I am used to figuring out the procedures step-by-step for a solution
AD-2: I usually try to find effective solutions for a problem
AD-3: I usually try to lay out the steps of a solution
AD-4: I usually try to figure out how to execute a solution for a problem
Generalization
GE-1: I tend to solve a new problem according to my experience
GE-2: I usually try to use a common way to solve different problems
GE-3: I usually think about how to apply a solution to other problems
GE-4: I usually try to apply a familiar solution for solving more problems
Evaluation
EV-1: I usually find a correct solution for a problem
EV-2: I usually think of the best solution for a problem
EV-3: I usually try to find the most effective solution for a problem
EV-4: I usually think of the fast solution for a problem

5.2 Applying the CTS Questionnaires to Healthcare Users

The CTS questionnaires introduced in the previous section were administered to 20 participants (11 females and 9 males) whose ages ranged from 21 to 58 (M = 31.3, SD = 10.84). All participants were potential users of our two EUD environments for the healthcare domain: 12 participants were pharmacists, 7 were enrolled in a Nursing B.Sc. course, and 1 was a medical doctor and also a university professor. The pharmacists were at different points in their careers, with an average of 3.54 years of work experience (SD = 2.71). Before administering the CTS questionnaires to participants, they were informed about the goal of the research and about the features offered by the two EUD environments, even though they had not been asked to use nor to evaluate them.

A summary of the overall distributions of results along the six scale values of the 5 CT dimensions are depicted in the violinplots of Figure 5.

As introduced in 5.1, we adopted an original approach to assessing the level of each of the five CT skills. Specifically, the responses were considered as ordinal values, and were codified into three classes: *Low* (for responses ranging from 1 to 2); *Mid* (for responses ranging from 3 to 4); *High* (for responses ranging from 5 to 6). The responses were then grouped by CT sub-scale, assuming that each of the questionnaire items in each sub-scale were independent and identically distributed (i.i.d.). This has allowed us to consider each item as a separate response with respect to another item of the same sub-scale. Instead of aggregating the scores into their mean

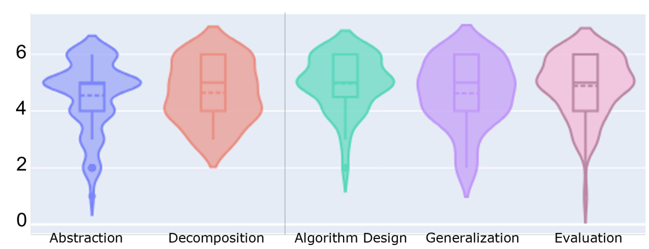

Figure 5: Violinplots of the CTS results, for each subscale of the CTS.

value, the raw scores were hence grouped into the three classification thresholds above. For each sub-scale, a χ^2 test was carried out to assess the statistical significance of the difference between the number of *Low*, *Mid*, and *High* responses of participants. For each sub-scale, the results were significantly different, at the significance level of .05, showing a level of skills for each sub-scale in the *High* range. The detailed test results are reported in Table 3, where, for each subscale, the absolute numbers of responses for the three classes are reported (*Low*, *Mid*, *High*), together with the χ^2 test result, and the p-value (p).

6 UNIFYING EUDABILITY WITH CT SKILLS

To unify users' CT skills with the EUDability of a EUD environment, the minimum requirement in terms of CT skills has to be derived from the results of the EUDability inspection-based evaluation.

Table 3: Results of the statistical analysis on the CTS responses.

	Low	*Mid*	*High*	Tot	χ^2	p
Abstraction	7	23	50	80	18.95	<.001
Decomposition	-	28	32	60	21.29	<.001
Algorithm Design	2	18	60	80	35.86	<.001
Generalization	6	28	46	80	18.32	<.001
Evaluation	1	22	57	80	35.36	<.001

When a EUDability dimension of a EUD environment is assessed as *High*, the minimum level of the CT skill related to that dimension that the user is required to possess is *Low*. On the contrary, a *High* level of CT skill is needed when a EUDability dimension is characterized by a *Low* degree. Finally, if a EUDability dimension is estimated as *Mid*, also the minimum CT skill level can be classified as *Mid*. In case a EUDability dimension is evaluated as *None*, meaning the system does not offer adequate EUD features, a system revision is mandatory and requires a further EUDability assessment.

Table 4 summarizes the EUDability evaluation of the environment supporting geriatric professionals in the creation of memory and attention exercises for older adults. The third column of the table also reports the minimum required level for each CT skill that users (geriatric professionals) must possess to effectively use the EUD environment.

In this particular case, it is crucial that users have high capabilities of Generalization and Evaluation, since the environment does not offer adequate features for component/exercise re-use and to test the results of the EUD activity.

Table 4: Required CT levels for the EUD environment devoted to geriatric professionals

EUDability	Level	Min. CT level	CT skill
Concreteness	*Mid*	*Mid*	Abstraction
Modularity	*High*	*Low*	Decomposition
Structuredness	*Midm*	*Mid*	Algorithm Design
Reusability	*Low*	*High*	Generalization
Testability	*Low*	*High*	Evaluation

Similarly, Table 5 summarizes the EUDability evaluation of the environment supporting pharmacists in the creation of robot tasks for galenic preparations. In this case, users do not need to reach an high level for any of the CT skills. The system could however be improved providing clearer instructions for robot task composition and simulation features (e.g., through a 3D visualization of the robot at work). Due to these limitations, the users must possess a medium level of Algorithm design and Evaluation skills.

According to the results of the administration of the CTS questionnaire to our sample population presented in Section 5.2, all the five CT skills of healthcare personnel reached the *High* level. The EUDability of the two systems reaches a demanding level of users' CT skills (*High*) only for the Reusability and Testability of the first tool, which meets the CT skill of all the participants in the study; therefore, we can derive that, from this preliminary and inspective study, the EUDability of both environments presented in this paper seems suitable to their target users.

Table 5: Required CT levels for the EUD environment devoted to pharmacists

EUDability	Level	Min. CT level	CT skill
Concreteness	*High*	*Low*	Abstraction
Modularity	*High*	*Low*	Decomposition
Structuredness	*Mid*	*Mid*	Algorithm Design
Reusability	*High*	*Low*	Generalization
Testability	*Mid*	*Mid*	Evaluation

7 LIMITATIONS OF THE WORK

Although the study involved participants who are expert in their respective fields of activity and were carefully instructed and controlled in the task of responding to the CTS questionnaire, we are aware that their number could limit the generalizability of our results. However, our major contribution lays in demonstrating how to unify the assessment of the EUDability of EUD environments by HCI experts with the assessment of domain experts' skills for CT. Having this goal in mind, the number of participants should not influence the validity of the approach.

On the other hand, any questionnaire or method carried out on its own is not sufficient to provide a robust methodology. There is always the need of an integration with a second method, possibly different (e.g., observing users performing tasks), in order to properly triangulate the outcomes of a user study. However, this kind of approaches are longer and costly. The CTS scale is a validated tool and our intent in applying it was also to demonstrate how it can be exploited for a faster and less expensive assessment of the human CT skills.

8 CONCLUSION

Our approach is a first attempt to unify Computational Thinking and EUDability through the combination of methods for assessing CT skills and EUDability dimensions. It allowed us to map the EUDability of two EUD environments from healthcare with the CT skills of domain experts, i.e., geriatric professionals and pharmacists. This mapping showed how the evaluation of a EUD environment may be done according to EUDability dimensions that mirror the users' performance in their daily work with EUD tools. The step-wise pattern is as follows: the assessment of the EUDability of a EUD tool (being it a prototype or a fully-fledged artifact); the assessment of users' CT skills; the unification of these two halves to assess the suitability (EUDability) of the EUD environment where domain expert work. Future work will systematize these steps into a methodology applicable to different domains. Despite some limitations outlined above, the current approach may remain valid in case of limited time and cost factors, e.g., in formative evaluations where organizations have limited resources.

ACKNOWLEDGMENTS

Partly funded by European Union Next-Generation EU "Characterizing and Measuring Visual Information Literacy" n. 2022JJ3PA5.

REFERENCES

[1] Jose Abdelnour-Nocera, Barbara Rita Barricelli, Arminda Lopes, Pedro Campos, and Torkil Clemmensen. 2015. Preface. *IFIP Advances in Information and Communication Technology* 468 (2015), V.

[2] Marcel Altendeitering and Sonja Schimmler. 2022. End-user Development for Smart Spaces: A Comparison of Block and Data-flow Programming. In *Proceedings of the 11th International Conference on Smart Cities and Green ICT Systems - Volume 1: SMARTGREENS,*. INSTICC, SciTePress, Setúbal, Portugal, 15–22. https://doi.org/10.5220/0010983200003203

[3] Thiago S Barcelos, Roberto Muñoz-Soto, Rodolfo Villarroel, Erick Merino, and Ismar Frango Silveira. 2018. Mathematics Learning through Computational Thinking Activities: A Systematic Literature Review. *J. Univers. Comput. Sci.* 24, 7 (2018), 815–845.

[4] Barbara Rita Barricelli, Fabio Cassano, Daniela Fogli, and Antonio Piccinno. 2019. End-user development, end-user programming and end-user software engineering: A systematic mapping study. *Journal of Systems and Software* 149 (2019), 101–137. https://doi.org/10.1016/j.jss.2018.11.041

[5] Barbara Rita Barricelli, Daniela Fogli, and Angela Locoro. 2023. Designing for a Sustainable Digital Transformation: The DEA Methodology. In *End-User Development*, Lucio Davide Spano, Albrecht Schmidt, Carmen Santoro, and Simone Stumpf (Eds.). Springer Nature Switzerland, Cham, 189–199.

[6] Barbara Rita Barricelli, Daniela Fogli, and Angela Locoro. 2023. EUDability: A new construct at the intersection of End-User Development and Computational Thinking. *Journal of Systems and Software* 195 (2023), 111516. https://doi.org/10.1016/j.jss.2022.111516

[7] Susanna Bergamaschi. 2008. *Demenza: 100 esercizi di stimolazione cognitiva.* Raffaello Cortina Editore, Milan.

[8] Marwa Boulakbech, Nizar Messai, Yacine Sam, and Thomas Devogele. 2023. Configuration approach for personalized travel mashup. *Concurrency and Computation: Practice and Experience* 35, 11 (2023), e6811. https://doi.org/10.1002/cpe.6811

[9] Karen Brennan and Mitchel Resnick. 2012. New frameworks for studying and assessing the development of computational thinking. In *Proceedings of the 2012 annual meeting of the American educational research association, Vancouver, Canada*, Vol. 1. AERA, Washington, DC, USA, 25.

[10] Henry Brodaty, Dimity Pond, Nicola M. Kemp, Georgina Luscombe, Louise Harding, Karen Berman, and Felicia A. Huppert. 2002. The GPCOG: A New Screening Test for Dementia Designed for General Practice. *Journal of the American Geriatrics Society* 50, 3 (2002), 530–534.

[11] Federico Cabitza, Angela Locoro, and Carlo Batini. 2015. A User Study to Assess the Situated Social Value of Open Data in Healthcare. *Procedia Computer Science* 64 (2015), 306 – 313. https://doi.org/10.1016/j.procs.2015.08.494

[12] Enrique Coronado, Dominique Deuff, Pamela Carreno-Medrano, Leimin Tian, Dana Kulić, Shanti Sumartojo, Fulvio Mastrogiovanni, and Gentiane Venture. 2021. Towards a Modular and Distributed End-User Development Framework for Human-Robot Interaction. *IEEE Access* 9 (2021), 12675–12692. https://doi.org/10.1109/ACCESS.2021.3051605

[13] Ingrid da Silva Carneiro and Ingrid Teixeira Monteiro. 2016. Communicability evaluation in systems created by designer-users. In *Proceedings of the 15th Brazilian Symposium on Human Factors in Computing Systems* (São Paulo, Brazil) *(IHC '16)*. Association for Computing Machinery, New York, NY, USA, Article 8, 10 pages. https://doi.org/10.1145/3033701.3033709

[14] Claudia de O. Melo, Jonathan Moraes, Marcelo Ferreira, and Rejane Figueiredo. 2017. A Method for Evaluating End-User Development Technologies. In *Proceedings of Americas Conference on Information Systems 2017*. AIS eLibrary, Copenhagen, Denmark, 1–10.

[15] Gerhard Fischer. 2002. Beyond "Couch Potatoes": From Consumers to Designers and Active Contributors. *First Monday* 7, 12 (Dec. 2002), 5 pages. https://doi.org/10.5210/fm.v7i12.1010

[16] Lauren B Flaherty, Allison Midden, and Benjamin T Mast. 2019. Psychometric evaluation of the symptoms of dementia screener (SDS) in a geriatric primary care population. *Clinical Gerontologist* 42, 5 (2019), 504–511.

[17] Giuseppe Ghiani, Fabio Paternò, Lucio Davide Spano, and Giuliano Pintori. 2016. An environment for End-User Development of Web mashups. *International Journal of Human-Computer Studies* 87 (2016), 38–64. https://doi.org/10.1016/j.ijhcs.2015.10.008

[18] Jonathan Grudin. 1991. CSCW: The Convergence of Two Development Contexts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New Orleans, Louisiana, USA) *(CHI '91)*. Association for Computing Machinery, New York, NY, USA, 91–97. https://doi.org/10.1145/108844.108858

[19] Ting-Chia Hsu, Shao-Chen Chang, and Yu-Ting Hung. 2018. How to learn and how to teach computational thinking: Suggestions based on a review of the literature. *Computers & Education* 126 (2018), 296–310.

[20] Till Janner, Robert Siebeck, Christoph Schroth, and Volker Hoyer. 2009. Patterns for Enterprise Mashups in B2B Collaborations to Foster Lightweight Composition and End User Development. In *2009 IEEE International Conference on Web Services*. IEEE, New York, USA, 976–983. https://doi.org/10.1109/ICWS.2009.46

[21] Eija Kaasinen, Susanna Aromaa, Antti Väätänen, Ville Mäkelä, Jaakko Hakulinen, Tuuli Keskinen, Joona Elo, Sanni Siltanen, Ville Rauhala, Iina Aaltonen, Juho Hella, Petri Honkamaa, Mikael Leppä, Antti Niemelä, Juha Parviainen, Markku Turunen, Jouni Törnqvist, Juha Valtonen, and Charles Woodward. 2019. Mobile Service Technician 4.0: Knowledge-Sharing Solutions for Industrial Field Maintenance. *Interaction Design and Architecture(s)* 2018, 38 (30 Jan. 2019), 6–27.

[22] Henry Lieberman, Fabio Paternò, and Volker Wulf. 2006. *End User Development (Human-Computer Interaction Series)*. Springer-Verlag, Berlin, Heidelberg.

[23] James Merisotis. 2020. *Human Work in the Age of Smart Machines.* RosettaBooks, New York, USA.

[24] Anders Mørch and Hege-René Hansen Åsand. 2006. Super Users and Local Developers: The Organization of End User Development in an Accounting Company. *Journal of organizational and end user computing* 18(4) (2006), 1–21.

[25] Anders I. Mørch, Valentina Caruso, and Melissa D. Hartley. 2017. *End-User Development and Learning in Second Life: The Evolving Artifacts Framework with Application.* Springer International Publishing, Cham, 333–358. https://doi.org/10.1007/978-3-319-60291-2_13

[26] Bonnie A. Nardi and James R. Miller. 1990. An Ethnographic Study of Distributed Problem Solving in Spreadsheet Development. In *Proceedings of the 1990 ACM Conference on Computer-Supported Cooperative Work* (Los Angeles, California, USA) *(CSCW '90)*. Association for Computing Machinery, New York, NY, USA, 197–208. https://doi.org/10.1145/99332.99355

[27] Fabio Paternò. 2013. End User Development: Survey of an Emerging Field for Empowering People. *ISRN Software Engineering* 2013, 532659 (2013), 11 pages. https://doi.org/10.1155/2013/532659

[28] Fabio Paternò and Volker Wulf (Eds.). 2017. *New Perspectives in End-User Development.* Springer, Cham. https://doi.org/10.1007/978-3-319-60291-2

[29] Daniel Rough and Aaron Quigley. 2019. Challenges of Traditional Usability Evaluation in End-User Development. In *End-User Development*, Alessio Malizia, Stefano Valtolina, Anders Morch, Alan Serrano, and Andrew Stratton (Eds.). Springer International Publishing, Cham, 1–17.

[30] W Joel Schneider and Kevin S McGrew. 2012. *The Cattell-Horn-Carroll model of intelligence.* The Guilford Press, New York, NY, USA, 99–144.

[31] Daniel Tetteroo and Panos Markopoulos. 2015. A Review of Research Methods in End User Development. In *End-User Development*, Paloma Díaz, Volkmar Pipek, Carmelo Ardito, Carlos Jensen, Ignacio Aedo, and Alexander Boden (Eds.). Springer International Publishing, Cham, 58–75.

[32] Christina Tikva and Efthimios Tambouris. 2021. Mapping computational thinking through programming in K-12 education: A conceptual model based on a systematic literature Review. *Computers & Education* 162 (2021), 104083.

[33] Meng-Jung Tsai, Jyh-Chong Liang, and Chung-Yuan Hsu. 2021. The computational thinking scale for computer literacy education. *Journal of Educational Computing Research* 59, 4 (2021), 579–602.

[34] Stefano Valtolina and Lorenzo Neri. 2021. Visual design of dialogue flows for conversational interfaces. *Behaviour & Information Technology* 40, 10 (2021), 1008–1023. https://doi.org/10.1080/0144929X.2021.1918249

[35] Changzhao Wang, Ji Shen, and Jie Chao. 2021. Integrating Computational Thinking in STEM Education: A Literature Review. *International Journal of Science and Mathematics Education* 20 (2021), 1–24. Issue 1949–1972.

[36] Jeannette M. Wing. 2006. Computational Thinking. *Commun. ACM* 49, 3 (mar 2006), 33–35. https://doi.org/10.1145/1118178.1118215

[37] Mustafa Yağcı. 2019. A valid and reliable tool for examining computational thinking skills. *Education and Information Technologies* 24, 1 (2019), 929–951.