SpiroPlay, a Suite of Breathing Games for Spirometry by Kids & Experts

Robby van Delden
r.w.vandelden@utwente.nl
HMI, University of Twente
Enschede, the Netherlands

Danny Plass-Oude Bos
Saxion University of Applied Sciences
Enschede, The Netherlands

Antje Jacoba Vivian (Vivianne) de With
Pediatrics Department,MST

Koen Vogel
Randy Klaassen
Nynke Zwart
HMI, University of Twente
Enschede, the Netherlands

Joyce Faber
Department of Pediatrics/Allergy Treatment Centre Deventer, Deventer Ziekenhuis
Deventer, The Netherlands

Boony Thio
Mattienne van der Kamp
Pediatrics Department, Medisch Spectrum Twente (MST)
Enschede, The Netherlands

ABSTRACT

We have built and implemented a set of metaphors for breathing games by involving children and experts. These games are made to facilitate prevention of asthma exacerbation via regular monitoring of children with asthma through spirometry at home. To instruct and trigger children to execute the (unsupervised) spirometry correctly, we have created interactive metaphors that respond in real-time to the child's inhalation and exhalation. Eleven metaphors have been developed in detail. Three metaphors have been fully implemented based on current guidelines for spirometry and were tested with 30 asthmatic children. Each includes multi-target incentives, responding to three different target values (inhalation, peak expiration, and complete exhalation). We postulate that the metaphors should use separate goals for these targets, have independent responses, and allow to also go beyond expected values for each of these targets. From the selected metaphors, most children preferred a dragon breathing fire and a soccer player kicking a ball into a goal as a metaphor; least liked were blowing seeds of a dandelion and applying lotion to a dog to grow its hair. Based on this project we discuss the potential and benefits of a suite-of-games approach: multiple games that each can be selected and adapted depending on personal capabilities and interests.

CCS CONCEPTS

• **Human-centered computing** → *Empirical studies in ubiquitous and mobile computing*; **User studies**.

KEYWORDS

Asthma; children; co-design; serious games; incentive; spirometry;suite of games; metaphors.

CHI PLAY '20, November 2–4, 2020, Virtual Event, Canada

ACM ISBN 978-1-4503-8074-4/20/11…$15.00
https://doi.org/10.1145/3410404.3414223

ACM Reference Format:
Robby van Delden, Danny Plass-Oude Bos, Antje Jacoba Vivian (Vivianne) de With, Koen Vogel, Randy Klaassen, Nynke Zwart, Joyce Faber, Boony Thio, and Mattienne van der Kamp. 2020. SpiroPlay, a Suite of Breathing Games for Spirometry by Kids & Experts. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '20), November 2–4, 2020, Virtual Event, Canada.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3410404.3414223

1 INTRODUCTION

1.1 Asthma and Spirometry

Asthma is one of the most common pediatric chronic diseases with a prevalence of 7-10% among Dutch children [3, 65]. It is associated with airway inflammation and airway obstruction, and characterized by symptoms such as wheezing, shortness of breath, chest tightness and coughing [3]. Childhood asthma imposes a significant social burden on the affected child and the family. It may prevent children from participating in play and recreational activities, may hamper social contacts, and lead to school absence and reduced school performance [13, 57]. As a result, children with asthma have a reduced quality of life (QoL) in comparison to their healthy peers [30, 50, 62].

The Dutch lung alliance states that regular monitoring of asthma control results in fewer symptoms and a better quality of life [36]. Spirometry provides physiologic parameters of the volume and flow of air, revealing the key features of asthma (i.e. airway narrowing) [32]. Typically this requires inhaling fully, and exhaling forcefully and completely into a tube connected to a measurement device. Performing one attempt of a sprirometry test in general for a child takes about ten seconds and less than eight attempts per session. The longitudinal data could enable healthcare professionals to get insight into the dynamics of the asthma status and optimize treatment, preferably measuring once or twice a day upto a year. Modern technology including low-cost turbine-based hand-held devices giving reliable data [10], can facilitate this, therefore we used NuvoAir's spirometry device.

However, a recent home monitoring study in the Netherlands with existing tools indicated a clinically significant margin for improvement: only 66% of the total spirometry measurements were technically correctly blown in comparison to the 92% in the hospital setting [19]. Spirometry measurements require maximal effort and patient cooperation to yield an accurate outcome measure. This can be achieved with professional guidance [19, 23, 31]. Moreover, many studies showed that, without guidance, children were not always compliant with the instructed number of spirometry measurements to perform in their daily life setting [4, 33].

The spirometry tests do have to be executed correctly to be comparable to previous measurements from the same person, or to compare the measurements to other (healthy) persons based on age, height, sex, and ethnicity using the Global Lung function Initiative (GLI) table [14]) or other databases including peak flow [21]. In our case, the target group includes children in the Netherlands, mostly Caucasian, from 6 to 12 years old. Children from this age range can reason logically thus allowing certain co-design methods [12].

1.2 Objective: Good Spirometry

To achieve proper execution of spirometry, according to Graham et al. [16], a spirometry test-procedure for children should at least contain: 1) a maximal inspiration, 2) a maximal forced expiration, and 3) expiration until completely exhaled. Furthermore, the right posture (i.e. sitting upright, chin slightly tilted upright, and feet on the floor) should be maintained.

In the context sketched above, we aim to improve unsupervised execution of spirometry tests for children and increase the likeability of these tests as a step to work towards long term compliance, by applying elements of game design.

1.3 Approach: Co-designed Metaphors Embedded in Games

Here we explain our starting point with accompanying elements of how we approach this project. In recent years, many research papers have built on the widespread application of gamification [34, 48], often making use of the trinity of *Points*, *Badges*, and *Leaderboards*. Based on our experience and interests so far, we start closer to a serious games perspective, namely persuasive games, where the focus is on developing an actual game rather than augmenting the task with game elements as explained by Deterding et al. [8]. We will refrain from labeling our game with the term *persuasive games* nor *serious games*, as persuasive can have negative connotations [58] and arguably seriousness has inner-contradictions and sets wrong expectations [49]. We do build on (game) design principles such as rewards and (lack of) punishments, affordance, juiciness, and game modes [49].

This also relates to fundamental research regarding dopamine release in games and positive effects of auditory cues and its possible effects on learning motoric behavior [18]. We aim for challenging fun game-like activities, including rewarding auditory cues with real-time visual feedback. We also follow Janssen et al. in their assumed expected benefits of bringing the expert (cf. therapist) closer to a game designer. To this end, we closely collaborate with two pediatric wings from the Deventer Ziekenhuis and the Medisch Spectrum Twente hospital, and many aspects of the games are based on their input. We also make use of the Persuasive System Design model [42]. Besides the metaphors, we include explanatory animations and feedback addressing reduction and tunneling. With our embedded goal setting we also apply tailoring and personalization.

1.3.1 Co-Designed Set of Metaphors with Children. We build on the philosophy of co-design and performed multiple co-design sessions with children [12] [1]. We envisioned that some of the games that the children would come up with might successfully trigger a first required behavior (e.g. full inhalation), whereas another succeeds in triggering other elements of behavior (e.g. hard expiration). We postulate that therefore our system will work better when multiple adaptable games are offered in a system, which allows offering a game fitting the wanted behavior for a specific person that has difficulty with this (e.g. giving the second game if the child repeatedly does not blow hard enough). Unfortunately, we know that validating this is extremely hard and will be beyond the scope of our research. Instead, we look at likeability, how we can generate several metaphors, and how children respond to such a game. We will use related work and parts of our results that do point to possibilities (e.g. children liking to play what they have not yet played), but we do not yet claim this to be validated.

1.4 Novelty and Added Value

The *novelty* lies in providing and creating incentives for spirometry addressing three targets of spirometry and starting this from a set of games inspired by children. We use a behavior-change support system perspective where we try to apply the power of physically related real-world metaphors in the games [6]. We will offer a selectable set of games fitting a person where each game is adaptable and uses tailored goal setting based on their capabilities, this is known as a *suite-of-games* approach [59] [2].

This work is part of a larger project in which we investigate whether providing an attractive co-designed game with appropriate representations (i.e. metaphors) and personalized feedback and feed-forward will positively impact the quality as well as the adherence of home spirometry in asthmatic children. To this end, we use NuvoAir's device in combination with the game running on a tablet. The outcomes related to quality (including a fault detection algorithm) are related to the medical domain and are beyond the scope of this paper. In a study in the hospital, we did evaluate the generated metaphors on likeability (i.e. in this case limited to preferred, non-preferred games, and explanation of what they liked and disliked). Aside from the games that were only based on children's input, a few metaphors are related to what is in use in current practice or what we found in the literature which will be described in the related work section.

The *added value* of this paper is twofold and relates to both process and application. One, in sharing the origin, our adaptations, and execution of our approach involving children and experts, we show the possibility and attainability of this approach even in a

[1] Following Stålberg et al. we believe that an adult adopting a child perspective cannot interpret the child's perspective accurately, not even when engaging with the same situations [53]. Although we use the term co-design, the involvement with only four afternoon sessions and a pilot is less than typically done in this field, we do (mis)use the term co-design, as we aim for giving children a real voice in the design process [12].Note, that keeping close to children's input resulted in not always the most (adult's) logical mappings of breathing to in-game actions, whereas Nacke et al. [35] point out the importance of this mapping in breathing or other biofeedback games, reiterated by [41, 67], and Tennent et al. [54] mention frustration of players when the game does not follow a particular affordance for mapping.

[2] See [41] for a non-tailored version of a 'suite of games' incorporating breathing as an input for people with Cystic Fibrosis (CF), and [2] for an extensive set of breathing games for CF designed with therapists and many developers.

health care setting and provide several lessons that can help solve future problems. Two, we created and share a set of metaphors, five important requirements (identified post-hoc) to facilitate metaphor development, and other lessons and considerations resulting from creating the metaphors. Relevance goes beyond the spirometry context. The process, findings regarding the application, and the suite-of-games approach might also be inspirational for mapping (unusual) input activity in general, including other breathing games.

2 RELATED WORK

Several commercial games include breathing as an input, including Nintendo DS' games that rely on microphone input to spin or propel objects, blow up balloons, or extinguish a fire [37, 38]. There are also commercial systems for respiratory muscle training for children with lung-related diseases [61, 68]. Furthermore, various studies on breathing games target Positive Expiratory Pressure (PEP) physiotherapy for people with Cystic Fibrosis (CF) [41, 67], relaxation exercises [1, 43, 51][3] including an attempt of calming children during drawing blood [52], and investigating entertainment-purposes [24, 35, 54].

There are however, two important differences with these studies and spirometry. Firstly, one attempt for a spirometry test takes a few seconds, where as PEP games takes longer (e.g. 5 min [2], 8 min [41], or even 20 min [67], 1-4 times a day [41, 67]). Secondly, spirometry tests include only one cycle of breathing, whereas rhythm, complex patterns, or frequency of breathing can play an interesting role in other types of breathing games [24, 41, 43, 51, 54, 67].

We continue with an overview of various spirometry-related studies. When looking specifically at home spirometry, there seems to be an added value of devices that measure several parameters (including total flow and flow in the first second) over older devices that just measured peak flow [33, 55]. Thompson et al. used measurements beyond peak-flow and looked at spirometry performed at home using a hand-held device with 67 children between 9 and 18 years-old [55]. The authors indicate that quality assurance was increased by showing correcting instructions for the next maneuvers based on the errors made during earlier attempts. The software awarded 500 'points' for three acceptable tests and 200 'points' if they tried six times but failed fulfilling the criteria of a proper test session. They found that the quality of the maneuvers was significantly lower for 9-12 year-old children compared to a group of 13-18 year-old children, which indicates the importance and challenge of measurement and instructions for especially younger children (c.f. [33, 40]).

Mortimer et al. [33] did a similar study with 92 children aged 6-11 years old and using the same device as Thompson et al. [55]. The authors used a similar point system. The points were now saved, and various prizes were offered in exchange. The authors compared the test to values from that person to their highest result from a previously acceptable test. Due to limitations of the device both Thompson et al. and Mortimer et al. could not yet provide feedback on the middle part of the flow-volume curves. Instead, they had to rely on visual inspection for that element of the tests.

[3] Note that although inspiring the study in [1] regarding 'The Journey to Wild Divine', was strongly critiqued on several points [69].

2.1 Incentives and Metaphors in Spirometry

Several commercial applications and studies do go beyond points as incentives. In the coming sections we will describe several examples that helped us in our design process and understanding of spirometry and could help other developers and researchers by providing a frame of reference and inspiration. Coates et al. present key considerations for performing spirometry. Children can be more willing with games such as a *'birthday cake with candles that extinguish during exhalation, a sailboat that moves across a pond, or a balloon that expends and pops when exhalation is complete. However, some technologists find this more distracting to the children than helpful.'*[7, p19]. Furthermore, they argue a real-time presentation of data is needed to obtain the best results by instant coaching, and feedback could then even allow the technician to early on abandon an incorrectly performed test. Vilozni et al. created SpiroGame, including two games with real-time feedback [64]. The first game focused on inhalation and exhalation, using a 30s breathing exercise showing a caterpillar going through an apple. The second game contains a bee that flies over a fence via flowers. Targets included quiet inhalation exhalation volume, total inhalation, peak flow of exhalation, and total volume. The caterpillar game was played as introduction and then either they played the bee game or a commercially available incentive of blowing out a candle. In the study the researchers noted that *'Incomplete expiration was common with the candle-blowing method (41 of 102 children), but rare with the SpiroGame (six of 102 children)'* [64, p2201]. Furthermore, only two of the 48 children playing the candle game exhaled longer than a second, one of the parameters often used. However, the game was played by very young children (3-6 years).

In a follow-up study Vilozni et al. used the same supervised multi-target game (i.e. games going beyond only peak flow) aiming for: full inhalation, instant forceful expiration, and long completed expiration [63]. In their system the target was also set based on personal data and was adjusted in 10% steps depending on whether the goal was reached or not. The 341 young children (2-6.5 years) were instructed by an 'experienced pulmonary technician'. The tasks being unclear within the game was the leading reason for unacceptable performance, in total resulting to only 22% rejected test session (triplicate spirometry). Vilozni et al. report a significant increase in peak flow as well as volume related values for their audiovisual multi-target game, when compared to standard verbal coaching [63]. The authors point out that more than just single targets are needed: *'games need rules, goals, and feedback to regulate the learning process'*. Furthermore, they also suggest using several games to each teach a different step in spirometry.

Other studies using single target games did not lead to sufficient correctly performed tests. For instance, Grachhi et al. used a single target of blowing out five candles and blowing up a balloon metaphor [15]. Such a target increased the reproducibility as well as (the input-related) height of values linked to instant forceful expiration. However, overall, it significantly reduced the overall performance of executing spirometry and total volume exhaled for their 88 participants of 4-8 years old. Instead using the same platform Nystad et al. did conclude that with a single target game reasonable acceptability and reproducibility can be reached when trained properly, for their 641 participants from 3-6 years overall

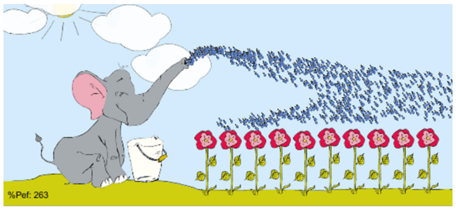

Figure 1: Picture of the elephant metaphor that waters flowers used by Gül [17], provided by Christian Dormeyer.

(but especially from 4 years on, e.g. 51% for 3 years old, whereas 78% for 6 years old) [40].

2.2 Recent Use of Multi-target Incentives

More recently, Gül in his PhD thesis shares a description of a device and accompanying games for real-time feedback mechanism and verification thereof focusing on children up to 6 years-old [17]. The system includes two standalone games focusing on practicing spirometry. The first is related to inhalation and exhalation where a bird is going up and down based on inhalation and exhalation, the bird flies forward constantly and catches mosquitoes. The other game is a good example of a multi-target incentive. Here an elephant has its trunk in a bucket of water and then waters a field of flower seeds, see Figure 1. The game targets four elements: breathing in rest, full inhalation, and forceful exhalation both long (fully exhaled) and forceful (peak flow). The water in the bucket first moves up and down based on breathing in rest. Then when fully inhaled the water is emptied from the bucket and the trunk moves up. While exhaling the water is sprayed over the field with seeds. Upon proper execution first the three furthest flowers emerge, which is based on the percentage of expected peak flow, and then the next seven, which are aligned with the percentage of expected volume reached.

Gül [17] and Dormeyer [9] performed tests with these games and their device with 22 children [4]. The first game was self-explanatory and tested with healthy children that all understood and had fun playing the game. The second game with the elephant was '*self-learning and interacted with the capabilities of the children*'. Many children could perform a reasonable spirometry with playing it once or twice (indicating 15 of the 22 [17]).

2.3 Commercially Available Metaphors

Several games already exist in commercial applications, where many still focus on peak flow, including those applied in the literature discussed before: blowing out candles, moving clouds, and popping balloons with an arrow [64].

One preferred digital metaphor applied in one of our hospitals is the milkshake metaphor. Where a child has to drink a glass filled with a milkshake through a straw (this is now part of the Vyaire products). Three other metaphors that can be chosen in the same software, are a kid blowing bubble gum, a wolf blowing away a house of straw, and a dragon with a marshmallow [5, 26]. CareFusion has another software-device combination which offers an incentive in the form of bowling, where full lengthy expiration is targeted, for an image see [11] and for a description see [22]. Vyaire's Vyntus and SentrySuite software offer a total of ten different animations including the earlier mentioned candles and bowling. Furthermore, there is a 'SeSBalloon', 'Speedometer', 'levitating octo-tube ball', 'Icarus in flight', 'flying toaster', dice, 'Hot Air Balloon', and 'Wood balls'. The settings seem to respond only to a value (FEV1) related to peak flow at the start [5].

Welch Allyn provides a metaphor in which a fireman with a (broken) hose is extinguishing a house on fire, which is targeting both peak flow and total volume of exhalation [44], for visualizations see [66]. '*The magnitude of and distance that the water flow extends to the house is directly related to achievement of the peak flow rate goal [..] The degree of the fire extinguished is directly related to achievement of the goal value for total volume expelled for the particular test subject once the water flow has reached the fire in the animation.*' [44] In the case the objectives are reached a rewarding message will be shown such as 'good job'.

Air Next (NuvoAir) is a handheld device that can be coupled to a smartphone. As an incentive there is a bar graph that shows the target volume value in the shape of a water column (including gas bubbles) to be filled and shows concurrent motivational messages such as 'Keep going!' [39].

The manufacturer MIR offers (at least) two different incentives. The first is for their handheld device and is similar to the water column incentive described above. In their system real-time instructional feedback is depicted (instead of motivational and performance feedback) including 'blow out faster' [29]. The second incentive is for their SpiroLab device with integrated printer and screen, it is the unveiling of a customizable picture or movie by moving a red 'theatre' curtain to the side [28, 47].

2.4 Takeaway and Connection to Related Work

Similar to Thompson et al. we also set children specific criteria primarily based on Miller et al. [31]. In our study we focus more on the idea that the test should also be entertaining. Kozlowska and Aurora even state: '*The biggest challenge is to make the testing fun.*' [22, p267]. Following Vilozni et al. we want to focus our games or 'visual incentives' on the several steps of the spirometry maneuver, where (different) metaphors can result in different executions, at least for younger children [22, 64]. Vilozni et al. indicate that the right metaphors might lead to significantly higher peak flow and volume [63]. Important steps are to make the games likeable, the phases of spirometry understandable in a way fitting the age group, and allow for adjustable targets to keep children entertained [22].

3 METAPHOR GENERATION & SELECTION

Our aim was to generate 100 metaphor ideas for spirometry in co-creation with children (and experts), to select the more promising ones and then specify further and implement a selection of these metaphors. Our first 28 metaphors were generated during a kick-off meeting with representatives of all the project partners with different backgrounds: including technical medicine, computer science, psychology, HCI, and a pediatric doctor. We first identified some of

[4] The explanation of these findings can be found in [9], quotes are translated by the first author combining [9, 17], we were unable to access the original academic work.

[5] The settings include difficulty level adjustments [25, 27] a full list of incentives, settings, and accompanying icons were made available to us [25].

the related work discussed in this paper. We then looked at popular activities and interests of children we knew, and added ten other ideas.

In preparation for the meeting and in parallel with the project, a high school teacher started a small HCI project with two of his pupils [6]. Their project also led to ten metaphor ideas, including blowing bubble gum until it pops, blowing leaves off a tree, a dragon spitting fire, and a dandelion. A few metaphors also had (unbeknownst to them) overlap with related work (e.g. a dragon, and blowing out a candle). They presented eight of the pictures and asked 26 pupils from the first-year class (11-13 years old) to rate these on a scale 1-10. This resulted in a preference towards blowing seeds of a dandelion (6.54), blowing a party horn (6.42), and a dragon spitting fire (6.15), others averaged 5.1 in range of 4.08-6.08.

Following co-design principles it is important to include the children from within the actual age range in several steps. Fitting the size of the research the involved children were assigned a so-called informer role. We did four co-design sessions with 6-8-year-olds. We then implemented three metaphors and performed a pilot test. Adaptions had to be made for the development phase for optimal implementation and execution. Finally, with a different set of end-users we did an end evaluation in a hospital setting which is discussed in the results section of the paper. Below we will describe how the co-design sessions build on Fails et al. [12].

3.1 Ethical Considerations of the Study

Participants of the co-design sessions did not get a monetary remuneration, instead they received a toy windmill and a small cookie after the second session. Importantly the activities themselves were intended to be joyful. We had a long-term yearly consent signed by both legal guardians for a number of studies with the children at the center. Instructions and templates regarding use of consent and studies were given, taught, and discussed with the high school teacher. We had two different types of ethical approval from our involved institutes. For the daycare design sessions and pilot study, including a broader way of working with the daycare center we gained ethical approval for this study from the faculty under RP 2019-14. For the evaluations at the hospital, including the likeability and understanding aspects from this study the national central committee on research with human subjects (CCMO) deemed this research not medically oriented and the research protocol was subsequently approved by the MST's local advisory board under KH19-17, and is registered in the Netherlands Trial Register (trial no. NL8048). It is important to note that our design sessions were not intended to, and not deemed fitting for, diagnosis of asthma, especially since pre-medication followed by post-medication testing is not done. We took special attention to communicate this carefully to the parents and children.

3.2 Participants

The sessions took place at a local out-of-school care. In total 12 children (8 boys, 4 girls) participated in the design sessions. We included only 6-8 year-olds, as the out-of-school care mainly provides care for younger children. This fits a large part of our target group.

The first two sessions were group-based and had one adult moderator. In total there were three adult participants, only the second session was done with two adults present.

3.2.1 *Experts.*

In order to filter the generated metaphors to the most promising set, we included input from spirometry experts via a survey. This included four lung function experts and three employees regularly responsible for spirometry at the pediatric wing of the MST [7].

3.3 Co-design Sessions & Expert Selection

In total we had four sessions at the local daycare center. The sessions are following three basic phases. A divergent phase, in which we generated many ideas. As this was a primary focus, we spent two sessions on this. A convergent phase, where we selected the most fitting ideas. And finally, a specification phase where instead of using requirements, we worked out what should happen for different errors of executing a spirometry test that could occur.

Each session was following several guidelines regarding setting the right environment, group size, getting to know each other, and rewards and attribution, as explained by Fails et al. [12].

We had a group of six children in the first two sessions. In the last two sessions, the adult participants worked with one or two of the children, as it was then more focused on choices and thus to prevent more peer pressure changing these choices. We used an activity room adjacent to the main room in the daycare center.

Building on the lessons and approach of Fails et al. the sessions were set up around the paradigm of:

- *Snack time* organized by the day care center, we only showed our faces, and told our names (taking about 15 min before the start).
- *Circle time* getting together and explaining both the bigger picture of the project and the goal of the session. The first two sessions took approximately 15 min, later only 2-5 min.
- *Idea generation* generating ideas or selecting and working out ideas, often using post-its and felt tip pens, where needed helping in writing down text. Duration varied per session, respectively about 30/25/25/40 min.
- *Round up* explaining what would happen with their work and what we would do in a next session. In the first two sessions this took about 15 min, in the last two about 5 min. After the second session they got to keep a toy windmill, after session 3 and 4 they got a small cookie.

To bring in the right atmosphere, in the first two sessions we started in a circle on the ground, where the three adult participants also sat on the ground. We did this combined with, for instance, triggering peer-explanations to facilitate working towards a feeling of equality instead of probing for answers. The first session had an icebreaker where children had to blow bubbles in a glass of water (related to one of the hospital's favorite incentives). In the second session they could blow in a toy windmill, that they got

[6]The benefit we foresaw for this project, was that younger people (16 to 18 years old) might have an easier time imagining how it was to be a child, and to translate their ideas. Something we could not investigate in our project in any detail.

[7]The experts were medical doctors with more than 5 years of experience. The other employees also included co-authors 3 and 9 of this paper having a technical medicine background.

Figure 2: Drawings made by children, some accompanied with text by children or adult participants. The drawings show the following metaphors that are selected (in adapted format) in the end: a fishing metaphor, penalty in soccer, a running match, and a car that gets more beautiful.

to keep afterwards. In the second session the children tried the actual spirometer showing comparable graphs upon completion (with a single-use turbine). Each week we mimicked breathing difficulties by breathing through a closed tightened fist [8]. In the next paragraphs we briefly go over other elements that are of importance on how we planned and executed each session.

3.3.1 Session 1 Divergent - Metaphor Generation. In session 1 we had three rounds of the idea generation. Round 1, *individual ideas*: we asked the children to draw an initial idea with felt-tip pens on post-its. Round 2, *inspiration cards*: we explained how something else can help to come to new ideas (an associative mindset). Therefore, we created six inspiration cards that contained a topic and drawing, including an animal (depicting a cat), superheroes (depicting a superman), and nature (showing a tree) among other depictions, which we loosely based on recurring topics from unrelated 'co-design' sessions, and we placed these at the center of the table. Round 3, *inspired by each other*: each child explained one of their own favorite ideas, and the other children were asked what they liked about it. If no feedback was given, an adult participant made a compliment and where needed helped to explain in case the child was shy. We then explained that all ideas would be gathered from the table on an idea board, that would be kept there for the three days up to session 2, see Figure 2.

3.3.2 Session 2 Divergent - Metaphor Generation. Session 2 was a continuation of the metaphor generation, done in a similar fashion as session 1 three days after. This session we exposed the children to actual spirometry, which was purposefully left out of the first session as not to affect idea generation from the beginning. All the children who wanted got a chance to use the spirometer themselves (each using their own new tube/turbine), which resulted already in some enthusiasm. We ended the circle time with a short reminder of a good spirometry blow: in, hard out, and out until empty. During the *idea generation* children were sitting around a table with two adult participants listening, prompting, and helping out three children each. We tried to do three similar rounds of idea generation, although this was less formally structured than the first session. Based on the last session, we decided to stimulate interaction between the children a bit more, as this helped to generate even more ideas and added more fun to the activity. When children had a hard time, we tried to trigger inspiration by asking about hobbies, what they liked, or what they did that day; here we also pointed at a new set of inspiration cards (in total we created 12). During the round up we added small heart-shaped stickers to mark the ideas they liked most.

3.3.3 Categorization and Filtering between Session 3 and 4. The first two sessions resulted in 56 recorded ideas: 38 in session 1, 18 in session 2; where especially those in session 1 were not all worked out in enough detail. We needed to filter the total set of metaphors to a graspable number of metaphors to choose from for individual evaluation. The total set also included the metaphors from the meeting, the high school student, existing applications, and related work. To include a spread of metaphors touching upon several popular topics, we started looking for recurring topics in the total collection of 100 (48+56 - overlap). For efficiency reasons, we did this with the three adult participants concurrently and co-located, where we majority-vote coded the metaphors.

We assigned topic labels in an iterative fashion. We stated a fitting category per metaphor, and then saw if there were similar but distinct or overlapping categories. For instance, we had several metaphors related to playing outside: blowing bubbles, going down a slide, and popping balloons with arrows. We tried to generate somewhat balanced groups but the group belonging to sports was the biggest. This made us split this category up in ball sports (8x) and non-ball sports (13x). Animals were also a popular topic (18, same animals were included in this count, 2x dog and 3x horses), followed by nature (9), vehicles (9), playing outside (8), and toys (6), people/jobs (6), things/objects (6), eating/drinking (6), technical abstract ideas (5, e.g. an image filter), buildings (3), and music (3). We picked four concepts from the six most filled categories. Where we discarded the more unsuitable concepts.

3.3.4 Session 3 Convergent - Individual Evaluation. With the 100 metaphors brought back to 24, session 3 was aimed at further selecting (to 12) and providing feedback on the generated metaphors, based on line drawings. As these steps required further work, this was performed 11 days after session 2. To prevent expected difficulties in rationalization ('why'), we took an indirect route asking what they would want to change, and using paired comparisons using only four metaphors at a time. Each session was done with one child and one adult. The three adult-child pairs were located in the same room using separated tables. Each adult participant conducted two sessions sequentially, resulting in the individual evaluation with six children in total.

For the selection, we used an adaptation of the *Fun Sorter* [45, 46]. We put four cut-out cards with drawings of a metaphor for the child. Each time this set of four was fitting one category (e.g. playing outside). We then explained briefly for each picture what would happen in the game during spirometry. Following ideas behind the Fun Sorter, the children had to put them in order of best to worst regarding three criteria, where we choose: Fun, Blowing properly,

[8]Do note this is not realistic and in current times with the COVID-19 virus should be reconsidered.

and Want to do again. To this end, we triplicated the cards. We stimulated the children to think out loud after each putting down or replacing of a card. We reiterated this for three categories, meaning each category would be labeled by at least two children.

3.3.5 Medical Expert Input Survey. We also incorporated opinions of the medical expert participants with the use of a survey. Adapted from the Fun Sorter, they had to sort four metaphors for each of the six categories regarding three measures: '*fun*', '*clear with exhaling*' and '*again*'.

Most experts responded between session 3 and 4 so we could take these remarks into account when selecting the most promising metaphors in the last session. Four of the seven experts did not understand or agree on the need to differentiate fun from what would be played again. Some additional remarks were given, for instance, to consider the pedagogical appropriateness of a metaphor. Feedback also included additional suggestions of how to specify the suggested metaphors.

3.3.6 Session 4 'Specification' Metaphor-related Feedback. For session 4 our goal was adding feedback regarding (other) issues that can occur during spirometry (these issues were selected from our experience based on Miller et al. [31] e.g. coughing during spirometry, or a delayed peak). As an ice breaker exercise we then let the children make small tubes from paper cups, which mimicked the turbine tubes with a realistically sized copy of the spirometer.

During *idea generation* we looked at each metaphor separately and went through the separate issues that could occur, using an accompanying drawing we explained what could happen. For each problem, we asked children what should happen in the game and to make a drawing or explanation on a post-it (in some cases this was also acted out).

3.4 Adaptations and Selected Metaphors

For the selection we looked at the ratings of children and the experts for preferences and triggering the right execution of spirometry. We worked towards two metaphors per category and aimed for including both the favorites of the children and the experts. However, later we decided to remove a playing outside metaphor where the expert preferred metaphor was the same as that for the children, whereas the second favorite only received two votes. In a similar way from the vehicle category we only picked the car, as this was the favorite of both, and the second favorite of the doctors received no votes of the children. Here the car was clearly preferred. Instead, we added another metaphor from the popular sports category, here the doctors were undecided over three metaphors. This resulted in a total of 11 metaphors, which are later described in Table 1 and for which the graphics are now ready.

In session 4 we generated in-game feedback for the different criteria of proper spirometry by Miller et al. [31] (e.g. preventing cough and glottic closure). However, with a team including the medical partners we found that this feedback could stimulate unwanted behavior, such as children on purpose triggering a flat belly landing for the springboard diving metaphor. After discussion with several of the lung function experts, we reduced the real-time feedback to only focus on full inhalation, forceful exhalation, and lengthy exhalation until completely empty, as emphasized throughout the other sections of the paper. Several of the metaphors were adjusted before implementation to more clearly address these three elements.

Due to time restrictions, only three were implemented completely for our test. We did work out eleven metaphors in drawings, based on input by children and experts, see Table 1.

4 SPIROPLAY APP: A SUITE OF GAMES

Building on the related work, a variety of applied practical principles, and the co-design sessions we made SpiroPlay. Similar to [52], SpiroPlay is a real-time responsive app created in Unity running on an Android tablet with a Bluetooth connection. We use Air Next (NuvoAir) a commercially available affordable turbine-based spirometer handheld device. The app uses a *suite-of-games* approach as it has select-able set of small games that can be adapted to the capabilities of the user [59]. Each game uses a multi-target incentive metaphor responding to full inhalation, hard expiration, and full exhalation.

Three finalized metaphors are included in the app and can be selected. These games are responsive in real-time to how the participant did regarding full inhalation, hard exhalation, and complete exhalation. All values are compared to a percentage of flow or volume with respect to an individually (for now manually) set goal based on personal parameters. This also allows the technical physician to set it to an adapted value depending on how their lung function is compared to that of an average child. A profile is saved in the app, where we keep track of the number of attempts per session (maximum of 8, less if 3 correct and 2 comparable) to not overload the child with the spirometry test. We save the key parameter as well as a log with the complete flow-volume curve. These curves are shown afterwards to allow to visually inspect these curves. Furthermore, textual feedback is given afterwards, where our algorithm provides additional instructions if spirometery was not performed well. Table 1 contains a short description of each selected metaphor. We will now describe the first three implemented metaphors in more detail [9].

The first implemented metaphor was a car, see Figure 3. At the start the car is standing still, based on full inhalation: a tachometer goes from left to right, the percentage of the expected peak then decides what car is depicted: getting a more beautiful car the better you do. The flow is mapped to the acceleration, and separately the achieved percentage of total expected volume decides where the finish is placed for '100%' but the car can go further.

The second implemented metaphor is popping balloons with a bow and arrow (building on the metaphor used in [64]). When inhaling the bow gets drawn where a bar in the top shows whether this is up to the expected inhalation volume. Based on whether the peak flow value was high enough the arrow will be set on fire. The reached percentage of volume of exhalation then decides how many of the shown balloons will be popped and how hard/high the arrow flies.

The third implemented metaphor is about diving from a springboard. Based on the percentage of expected inhalation a figure is walking to an indicated spot on the springboard. On reaching the wanted peak the figure then gets a cape. On reaching the expected

[9]We also refer the reader to our Github [60], it will include the art and code of the Unity project with fault detection after our work is published regarding those algorithms.

Name	Topic	Favorite	Start	Full inhale	Expiration hard	Expiration long
popping balloons	playing outside	6-8, experts	a bow is shown, and a row of balloons	bow is drawn	arrow is on fire	more balloons pop
car	vehicles	6-8, experts	slow car	tachometer, exhaust clouds	nicer car, turbo particles	progress bar speeds up, reaches finish
diving	other sports	6-8	at start of springboard	walks to end	flies up receives a cape	goes through clouds reaches the sun
hurdles	other sports	experts	athlete is standing	gets into pose (a silhouette)	jumps over hurdle (not tripping it)	reaches the finish
fishing	other sports	experts	side view person, on a peer with water	fishing net drops into the depth	a beautiful fish appears in the net	net is filled with fish
soccer penalty	ball sports	experts, 6-8	side view player ready a goal on other side	leg swings backward exaggerated, progress bar	ball gets particles (wind+fire), cheering	reaching the goal
disco bowling	ball sports	6-8	bowling alley side view	arm swings backward colour coded progress bar	disco ball out ceiling	ball rolling on lane, stripe on lane as goal
dog hairs	animals	6-8	naked dog on podium with a lotion bottle	dog gets foamed with hair growing lotion	spotlights turn on	the hair grows up to the floor
dragon	animals	experts, high school	dragon next to oil barrel at the far end wood	barrel is transparent, all oil is sucked up	fire changes color	fire reaches further campfire is ignited
dandelion	nature	6-8, high school	yellow dandelion with a lady bug on its stem	flower changes in seeds lady bug at top of stem	lady bug flies away	the last seeds are dispersed into the air
tides (beach)	nature	experts	a beach with one person	water retracts treasure chest sticks out	dolphin jumps out sea,	water rises and reaches feet, person runs away

Table 1: An overview of the resulting metaphors, most originated from co-design sessions. Shooting balloons builds on a metaphor in [63], the bowling is an adaption of the commercial version. The dandelion and the dragon (latter mentioned in a patent [20]) were generated by the high school student. Three metaphors have animations that finish after completion. For example, for the diving doing tricks and reaching the water with applause.

percentage of volume the diver reaches the sun and might even go beyond reaching that. After detected end of breath the animation finishes with the diver entering the water.

5 (PILOT) TESTS OF 2 METAPHORS

To verify the functioning and overall usability we performed a pilot at the daycare center. During this pilot, five children played with the first two implemented metaphors (car and balloon popping). We gave minimal accompanying instruction, showed the worked-out images of the metaphors, asked for feedback, and asked about their interpretations and what they liked. This pilot took a little less than an hour and showed severe technical issues with the connection.

Feedback on the other metaphors was gathered based on just their drawings. This way we got a tip to better align and depict the stages in the dandelion metaphor. The children also needed more detailed instruction about the whole procedure (e.g. sitting up right) and about the steps during spirometry (e.g. when to put the device in their mouth). In several subsequent technical tests we improved the performance and addressed these issues and feedback.

6 TESTING 3 METAPHORS IN THE HOSPITAL

To test the metaphors with a fitting target group of participants we performed a study at the pediatric wing of the Medisch Spectrum Twente hospital [10]. In this paper we look at preference of the presented metaphors and feedback on them.

[10]The study also had another attempt with a black screen, we added an instructional video, automatic detection of errors, and accompanying feed-forward and feedback, we also looked into the resulting values and type of errors.

6.1 Procedure and Measurements

The children performed a spirometry test with one of the three metaphors, equally distributed over the population. Children were participating voluntarily and could stop at any time. For each spirometry test children had to perform a minimum of three maneuvers and a maximum of eight, depending on the acceptability and repeatability of the values and execution. The technical physician observed execution and wrote down any feedback the child gave. Regarding usability, we also looked if the children took initiative to navigate through the app to select a certain game or skip a step, and observed what happened.

We then showed them a recorded gif of the other two games they did not play and explained the eleven metaphors using three pictures for each with an accompanying spoken-out scripted description. After the tests, the children had to fill in a questionnaire, which included questions about what happened in the game they played, what they liked most about the game, and what the least. Furthermore, we asked them for suggestions to improve the game even further. Inspired by the Smileyometer of Read et al. [45] we asked them to rate the three implemented games on a Visual Analogue Scale from 0-10 (instead of 1-5); 0 anchored as ‘Very stupid’, 2 ‘stupid’, 4 ‘not so fun’, 6 ‘a bit fun’, 8 ‘fun’, and 10 ‘Super fun!’ with accompanying smileys ranging from red to orange, to yellow, to green. We asked which three games they would prefer to play, and which three they would like to play the least.

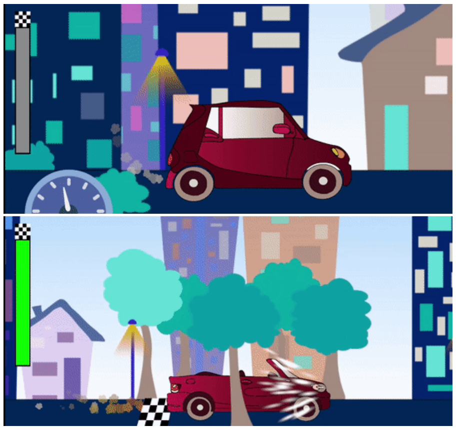

Figure 3: The car metaphor, above: showing the begin while inhaling (tachometer half way), and end when reaching the peak (car type) and just over the expected volume (finish).

Figure 4: On top, a depiction of the dragon metaphor, showing the dragon that emptied a barrel setting wood on fire. Below, the soccer penalty metaphor, showing a player that swung its leg to kick a ball and a ball in the goal.

6.2 Participants Hospital Study

The participants were children with pediatrician-diagnosed asthma in the age group of 6 – 11 years old. We had exclusion criteria for children with acute asthma symptoms, and children with comorbidities that could have additional difficulties with the instruction or execution (i.e. autism, developmental impairment, neuromuscular disorders). In total 30 children were included, 15 in the age range 6-8 years old, and 15 in the age range 9-11 years old, 9 of children were girls and 21 boys. We want to, in-line with Schell's suggestions, point out the importance of preventing a pre-defined sex-based stereotyping of preferences but also to consider making unlabeled use when differences arise to accompany preferences fitting the gender spectrum [49], so we do report this in the table.

6.3 Control and Understanding

Most children seemed to understand how the app works. Children showed clear initiative and seemed to understand the workings of the user interface (17x), a few (mostly timid impression) needed active encouragement but would not act out of own volition (3x), and the others just seemed to understand how it worked. About ten children already during the test asked whether they could, or actually did, click on next or skipped instructions.

For the balloon metaphor the physics in the game were not programmed well. In many cases the arrow did not go through the balloons although the balloons were popped. This was a main point of critique by children and in some cases even led children to stop the test prematurely.

Overall the children remembered and explained well what happened in the game. However, in their explanation they either emphasized the response to blowing hard or blowing long and none of the children emphasized both. Instead, they talked about what happened when inhaling and exhaling.

6.4 Parts Preferred, Disliked, and Suggestions

Many children could not come up with any negative points about the game (10x), and simply stated: *'I actually liked everything!'* or mentioned in their explanation *'it was more fun than a normal blow test'*. Two children instead stated they did not like the game they played. One even called the diving game *'boring!'* Ten children did not like that they 'lost': that they did not make it to rewarding visualizations (i.e. car-type, cape, fire), especially regarding the car metaphor (6x). One child remarked: *I had also wanted to get the car with the open roof!* Several suggestions for improvement also mentioned this: *'also stating that you will not make it to the quickest car, so you don't get disappointed'*.

Other suggestions included simple adjustments of changing the background into something livelier for the balloon and diving. For the car metaphor additional city elements were suggested such as a bridge or the Eiffel tower, and to have the car grow. More extensive suggestions were also given, often related to commercial games including shooters, angry-birds, or adaptable (point-)systems, for instance, to use a catapult and a ball with spikes instead of a bow and arrow. Furthermore, regarding the car, suggestions were made to get a fire boost, to get a new car, motocross motorcycle, or that instead a unicorn with sparkles could be chosen.

6.5 Likeability of Metaphors

Overall, the three implemented games received high ratings, averaging 7.8, 7.4, and 7.9 on a scale of zero to ten, for the car, balloon, and diving metaphor respectively. Remarkably, the highest rated metaphor (diving) also received three times a 0 rating, all of children playing that game, whereas no other game received this rating. Whereas the 30 children gave a 10 rating six, four, and ten times, for the car, balloon, and diving metaphor respectively.

When grouping all results the games that were preferred most to be played, where dragon (20x) and soccer (17x), note that three preferences per person means out of a total of 90 ratings with a 30 maximum, see Figure 4. The games that were most *undesirable* were the dog ($13\frac{3}{4}$x) and dandelion ($12\frac{3}{4}$x), here one child was determined to pick four, so we normalized to $\frac{3}{4}$pt ratings for this child's choices.

See Table 2 for the full overview of results. When looking specifically at the difference between two age groups, there are indications of some small differences. Younger children clearly prefer the dragon (and diving to some extent), and older children seem to prefer the tides and bowling more.

top 3	6-8 P	6-8 D	9-11 P	9-11 D	F P	F D	M P	M D	T P	T D
Car	2	2	2	5	0	4	4	3	4	7
Balloons	2	2	2	5	1	2	3	5	4	7
Bowling	3	3	7	0	3	1	7	2	10	3
Dragon	12	0	8	3	6	1	14	2	20	3
Dog hairs	2	$8\frac{3}{4}$	3	5	2	4	3	$9\frac{3}{4}$	5	$13\frac{3}{4}$
Hurdling	1	4	2	6	1	3	2	7	3	10
Dandelion	1	$6\frac{3}{4}$	1	6	2	4	0	$8\frac{3}{4}$	2	$12\frac{3}{4}$
Diving	6	$2\frac{3}{4}$	2	4	2	1	6	$5\frac{3}{4}$	8	$6\frac{3}{4}$
Tides	3	$4\frac{3}{4}$	7	3	6	0	4	$7\frac{3}{4}$	10	$7\frac{3}{4}$
Fishing	4	8	3	3	0	4	7	7	7	11
Soccer	9	3	8	5	4	3	13	5	17	8
rate 0-10	avg		avg		avg		avg		avg	
Car	7,9		7,6		7,5		7,9		7,8	
Balloon	7,9		6,8		7,8		7,2		7,4	
Diving	8,1		7,6		9,2		7,3		7,9	

Table 2: Table showing the three ratings for preferred (P) and disliked (D, i.e. liked least), as well as the rating of the three implemented games. The values are also grouped regarding age (i.e. 6-8 year old, and 9-11) and sex (i.e. female (F) $n = 9$, and male (M) $n = 21$).

7 DISCUSSION

Overall, the co-design sessions resulted in many interesting metaphors. We succeeded in using multi-target incentives and we generated 11 seemingly suitable metaphors, from which we currently completed and tested three that received a reasonable rating regarding fun. We do advise the reader to not take the preferences at face-value and simply work-out the preferred metaphors, as Read and MacFarlane mentioned also regarding evaluations with children '*information gained from a single group of children in a single place is not likely to be especially generalisable'* [46, p86]. Instead we encourage any interested party to take special consideration of the elements we address in the paper, especially those that follow in this discussion.

7.1 Likeability/Fun, Again - Again, and Ratings

The children were positive about the games in general. Younger children (6-8) rated the three games slightly higher on average (M=8.0, SD= 2.0) than the older children (9-11) rated them (M= 7.4, SD=1.9), which fits the finding of higher ratings found for children 8-9 versus 9-10 year-olds by Read et al. [45]. However, even based on a (fast approximation i.e. likely breaking the interval variable assumption) post-hoc analysis using a one-tailed independent t-test we do not see a significant effect of averaged rating between groups $t(28) = 0.86, p > 0.05$.

7.2 Fun vs Prefer to Play, or Not the Same?

Interestingly, the questions rating the three games (0-10) and selecting three games that they would prefer to play from the full set of 11 metaphors do not align with what one might expect. Of the 20 times that a game was rated with a maximum score of 10, only 14 times the accompanying metaphor was also selected from the full set. One reason could be our envisioned difference between games that would be *'Fun'* versus those they would *'Want to do again'*. We did anticipate this in the adapted use of the Fun Sorter in our co-design sessions. There as well, the children showed a difference in preference for these two criteria (fun vs again) in 15 out of the 18 cases. The children's explanations while making this choice also seemed to indicate that they imagined having different preferences if they would have to play repeatedly, unlike the most experts.

Read et al. show wanting to play a game again (i.e. comparative Again - Again table, with several systems yes/maybe/no) is strongly correlated with the results of the Smileyometer [45]. However, it would be interesting to further investigate whether there might still be a difference between a game that is fun to play and those that would be chosen to play (again) in a setting like ours. This in itself could be an influence on our results, as children who already performed the test with one metaphor up to eight times, might be negatively prompted for the selection of playing that game again from the 11 metaphors. This could point towards children liking diversity, wanting to play different kinds of games in subsequent sessions. Similarly, the large difference in ratings for some metaphors between children could indicate that having multiple metaphors fitting different children does have an important potential. A suite-of-games approach could address both issues.

To investigate this further, we compared the average rating for the game the children played versus the two other implemented games that they only saw animated gifs of. We conducted another fast approximation post-hoc analyses where we averaged the two games they did not play and did a paired two-tailed t-test. To indicate that this deserves further investigation, we present this informal result pointing towards a significant effect ($t(29) = 2.05$, $p = 0.05$) of children rating the game they played lower ($M = 7.1$, $SD = 2.9$) than they on average rate the other two implemented games ($M = 8.0$, $SD = 1.8$). However, an alternative explanation here could be that the expectations of the game are higher than the actual game play (e.g. the arrow missing balloons).

7.3 Real-world Link of Metaphors

Looking back at the three games that have been implemented currently, we see a certain lack of physical correlation. This is partially due to keeping close to the children's input while incorporating expert-inspired and development-based adjustments mostly later on in the process. This especially breaks the first two steps of proper execution of Chow's four-step suggestion for designing blended metaphors in persuasive systems [6, p52]. This might be one reason why none of the children clearly had an explanation of the game that distinguished long and hard. In several studies [35, 41, 54, 67] the importance of real-world related mapping, challenges to do so, and issues with real-world mapping in various breathing games are discussed. In a follow-up long-term study we want to investigate this further and see if the set of games with perceived stronger ties to real-world physics are also leading to better spirometry results than those with more abstract relations.

There are games which seem to have a reasonable 'embodied' fit, such as the tides metaphor, where children draw the water towards them by breathing in and then blow it towards the girl standing on the beach by expiration. One participant actually unprompted noted this also as a positive point of the diving metaphor: *'that* *the game moved with your feeling, so when exhaling then she went up'.* Analytically, the same seems to hold to some extent for the dandelion and dragon. Less of a fit can be seen in the car game, where inhaling gets a car to start (full rpm in neutral gear). In related work we have also seen both types of metaphors, including a caterpillar walking versus blowing out a candle [64]. The link of the metaphor itself to the real-world fit might also be of influence for the type of errors seen. The blowing out a candle was not only responsive to the measured peak by the capabilities of the system, but might also steer more towards generating a peak flow than complete exhalation, based on its real-world link.

7.4 Reaching Metaphor Goals

Kozlowska and Aurora make an understandable argumentation for how in-game goals should fit a personally adjustable target [22]. Using the bowling lane metaphor they state one should have the incentive animation such that when peak flow is reached the ball should be about four-fifth towards the skittles, and should then reach the end when 'completely empty' (residual volume). However, they also state that the skittles should not be knocked down, but that it should feel as if the child could reach it when trying again. This should in turn prevent stopping exhalation when the target is achieved. They suggest that when the tester is already satisfied the goal can be set lower to reward the child. In our project we choose to use the predicted values, an attainable target for this target group. We also believe that once a reliable maximum reproducible score is reached it is inappropriate or perhaps even unethical (e.g. unneeded deception), if this would not lead in achieving the in-game goal.

In our experience the completed animation is possible with a small delay or by providing targets above 100%. With this we think premature stop can be prevented. This can thus also follow the actual performance of complete flow (cf Vyaire's Viasys is seemingly responsive to the peak measured at start but shown at end of breath [56]). We postulate that for optimal results, this should be combined with goals that are first personalized (e.g. in % of expected values and choice of metaphor) and then tailored (e.g. based on values from a successful test). Following Vilozni et al. we urge to go beyond just peak-flow, and include concurrent goals for at least full inhalation, volume of complete exhalation, and instant forceful exhalation at the start. Where regarding the latter values, in this research we used peak flow as the maximum rate based on [21], but others might, considering age differences, use the volume until .5/.75/1s [22].

Over 100%. To further stimulate children, even when reaching the 100% of estimated volume goal (full exhalation) several metaphors allow to depict responses for these additional percentages. For instance, the car drives farther over the finish and the diver flies over the sun and beyond and for the fishing metaphor the fishes will bounce out of the net. When generating the metaphors, we did not yet realize the added benefit of this, so not all of the chosen metaphors have this opportunity. We do advise others to take this into account when generating additional metaphors or when using our metaphors.

Beyond binary. Not all metaphors respond gradually to blowing hard enough. For instance, for the diving metaphor the athlete gets a cape if the expected peak flow is reached, when the child does not blow hard enough this will not be shown. This makes it closer to a punishment than a reward, the child knows what is possible but does not achieve this. Alternatively, the car metaphor already provides positive reinforcement when part of the goal percentage is reached, as the car is upgraded to nicer models on the way to the target. This feedback does need serious consideration as in the results we also showed that children disliked (10x) that they did not achieve the completed animation, somehow especially for the car metaphor (6x).

7.5 Development Adjustments and Post-hoc Requirements

After the first revision based on the input from the involved medical experts, some metaphors were further adjusted by the development team. For example, when it required time-intensive visualizations and physics non-essential for the core idea of that metaphor (e.g. that the dragon would be flying).

During the co-design process we did not yet have access to a real-time responsive stream of NuvoAir's device. This was one of the reasons why the participants present during the co-design sessions did not realize the actual speed of the procedure in-game (researchers and children alike). The development team realized this once we gained access to the real data stream. This led to new requirements, and several metaphors had to be worked out differently while keeping close to the input on which they were based on.

7.5.1 Requirements for Multi-target Incentive Visualizations. The following five elements were introduced as important requirements during development and adjustments after the co-design sessions.

(1) *The full exhalation is depicted in an independent way from the inhaled volume.* We should be able to animate both when children did not inhale fully but do exhale fully, and when they inhale fully but stop prematurely.
(2) *Achieving full exhalation is depicted in a separate way from achieving the hard exhalation peak.* For some metaphors, these two things were mixed (making use of the same physics), especially because children in the first two sessions seemed to focus on inhalation and exhalation only. To at least be able to finish a somewhat imperfect spirometry test to their best capabilities this should not be dependent on each other.
(3) *No animation between inhalation and exhalation.* The exhalation should directly follow exhalation. Also, as we can make an estimate of the (remaining) volume however timing is harder, we were unable to predict in real-time how long inhalation or exhalation will take. We had no data for this parameter.
(4) *Both inhalation and complete exhalation indicate a clear goal that children want to reach.* Several metaphors did not clearly indicate how much was expected, thus not giving the right incentive to reach the expected values. Several metaphors now have a progress bar based on the expected values.
(5) *The metaphor is responsive for values going beyond 100% both for inhalation and exhalation volume.* This is explained above and in our vision could also address issues pointed out in the current use of some commercial systems.

7.6 Beyond Spirometry

The current application is fully oriented at spirometry, even more specifically for monitoring children with asthma. However, parts of the design process, the suite-of-games approach, and how we included the children might be inspirational beyond this context.

Breathing is seen as an interesting game control mechanism, even within the area of biofeedback, as it is both voluntary and involuntary, and is one-dimensional analogue but games can also use derivatives such as frequency of breathing [54], breathing patterns, or volume [24]. So it is not surprising that studies have looked into breathing as a direct control to enhance game interaction [24, 35, 54], to train relaxing breathing exercises [43, 51], or for positive expiratory pressure (PEP) exercises to release mucus [2, 41, 67].

Perhaps the generalizations of the post-hoc requirements we found (e.g. related to expected values (4,5), separation of values (1,2), and preventing animation in-between inhaling and exhaling (3)) would also be beneficial in such a context. In future research where one would also look more into the feedback and instructions, more focus could be given to design considerations based on other breathing game contexts. Marshall et al. [24] developed an automated breath-controlled bucking bronco ride leading to design considerations mostly related to tactics and feedback (loops) of the ride influencing breathing (and breathing the ride). Patibanda et al. [43] worked on LifeTree: a fully immersive breathing exercise, they found different kinds of design considerations, for instance related to subtleness of feedback and supporting self-awareness of breathing. Furthermore, such future work could also looking at the importance of the fit of physical elements of the hardware to the context and game [43, 51, 54], including emphasizing the user's own breathing sounds [43].

7.7 Personalized Unlocking in a Suite of Games

We designed a reward system, which rewards compliance of the child with unlocking new metaphors. This is related to the suggestion to reward the child by letting them choose to play a certain metaphor of choice once reliable spirometry tests have been performed [22].

Besides adaptable and adaptive goals within a game, we will also make use of the suite-of-games approach to personalize what game should be played considering the child's difficulties in their tests. For instance, if we see over time a child especially has difficulty with hard expiration, and if we would know that in general the car metaphor game is excellent in stimulating this quality, we can suggest this game more often. We will incorporate this by unlocking a personalized set of games after a set number of spirometry measurements. We will also unlock other metaphors based on reaching a preferred interval of testing.

The unlocking elements are inspired by the game mechanic of unlocking achievements and new levels related to goal achievement and more specifically to Zhao et al., who showed that feature release over time led to higher compliance for an exergame than offering all functionality at the start [70]. We are looking forward to completing a longer term test with children offering adaptive unlocking mechanics looking into the compliance and quality over time.

8 CONCLUSION

We have created SpiroPlay, an app to track lung function of children with asthma. The app includes three implemented games that run on an Android tablet connected to the Air Next (NuvoAir). Children could use it to try and perform spirometry. We used a *suite-of-games* approach, multiple small games (11 drawn out) from which one can be selected to be used and each game can be adapted to the capabilities of the user. We used and emphasized the importance of a multi-target incentive metaphor responding to full inhalation, hard expiration, and full exhalation.

To do this, we looked into related work, applied methods from co-design using both input from children gathered in four sessions as well as from medical experts, which resulted in a set of over 100 metaphors. From a selection of 11 of these metaphors, children especially wanted to play a dragon and soccer metaphor. Our application of the co-design methods showed it is attainable to also use this in a health care (research) project. We reasoned and discussed how it can be beneficial to use these elements of a *suite-of-games* approach (personalize the selection of a game and to adapt the game), where we found several indications that this approach might be fitting. Children often wanted to play what they did not yet do. Children can differ in what they liked, and based on current practices and related work regarding asthma monitoring, we see that goals should be adaptable to personal capabilities.

Besides sharing our approach, we also learned several lessons. We created a set of requirements and described points of attention to help the development of future use of metaphors for spirometry. We propose five requirements that, besides including clear goals, also state to clearly distinguish elements in the animation as recognizable implementable independent feedback responding to the user's three actions (complete inhalation, forceful exhalation, and complete exhalation). We also offer suggestions for how games can be unlocked over time based on the behavior it triggers, and a suggestion to provide clear attainable goals with visualizations for over 100% of predicted values. In providing these lessons and metaphors, we intend others to benefit during their research and development of future systems and to get more people to make use of it. So, we might not only help children to have more likeable games, but even make the life of a child with asthma itself a bit more likeable.

ACKNOWLEDGMENTS

This study was financially supported by an unrestricted research grant from the Pioneers in Health Care Innovation Fund, established by the University of Twente, Saxion, MST, ZGT and Deventer Hospital. We like to thank Monique Gorissen, Alejandro Moreno, and Floor Sieverink for sharing their input in numerous meetings and writing the proposal. Bridgette Connell for extensive proofreading. All the student assistants (Sil Tijssen, Rens van de Werf, Sterre van Arum, and Kevin Smid) included in the development. Cristina Zaga for improving the co-design sessions, Dennis Reidsma and Edwin Dertien for inspiring us with their work in breathing games. Iris Heerlien and Tamara Droogsma for their related thesis work. As well as Melvin and Noor and their teacher Job Burema. Finally, Myrthe Asbroek for her key role in facilitating sessions at the daycare center.

REFERENCES

[1] Krestina L Amon and Andrew Campbell. 2008. Can Children with AD/HD Learn Relaxation and Breathing Techniques through Biofeedback Video Games?. *Australian Journal of Educational & Developmental Psychology* 8 (2008), 72–84.

[2] Fabio Balli. 2018. Developing digital games to address airway clearance therapy in children with cystic fibrosis: participatory design process. *JMIR serious games* 6, 4 (2018), e18.

[3] Eric D. Bateman, Suzanne S Hurd, Peter J Barnes, Jean Bousquet, Jeffrey M Drazen, Mark FitzGerald, Peter Gibson, Ken Ohta, Paul O'Byrne, Soren Erik Pedersen, et al. 2008. Global strategy for asthma management and prevention: GINA executive summary. *European Respiratory Journal* 31, 1 (2008), 143–178.

[4] Patricia V Burkhart, Mary Kay Rayens, Marsha G Oakley, Demetrius A Abshire, and Mei Zhang. 2007. Testing an intervention to promote children's adherence to asthma self-management. *Journal of Nursing Scholarship* 39, 2 (2007), 133–140. https://doi.org/10.1111/j.1547-5069.2007.00158.x

[5] CareFusion. 2015. Spirometry PC Software. Retrieved March 19, 2020 from www.carefusion.com/documents/brochures/respiratory-care/cardiopulmonary/RC_SPCS_BR_EN.pdf

[6] Kenny K. N. Chow. 2019. Designing Representations of Behavioral Data with Blended Causality: An Approach to Interventions for Lifestyle Habits. In *Persuasive Technology: Development of Persuasive and Behavior Change Support Systems*. Springer, 52–64. https://doi.org/10.1007/978-3-030-17287-9_5

[7] Allan L Coates, Brian L Graham, Robin G McFadden, Colm McParland, Dilshad Moosa, Steeve Provencher, et al. 2013. Spirometry in primary care. *Canadian respiratory journal* 20, 1 (2013), 13–22.

[8] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. From Game Design Elements to Gamefulness: Defining "Gamification". In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (Tampere, Finland) *(MindTrek '11)*. ACM, New York, NY, USA, 9–15. https://doi.org/10.1145/2181037.2181040

[9] Christian Dormeyer. 2015. *Animationen – spielerisch früher zur erfolgreichen Messung?* Technical Report. Starnberg, Germany. KongressReport GPP-Kongress, https://www.pari.com/fileadmin/user_upload/PARI.com_SpiroSense/Docs/120D0006_SpiroSense_Kongress_Report_GPP_2015.pdf.

[10] Elsabe Du Plessis, Frederiek Swart, D Maree, J Heydenreich, J Van Heerden, TM Esterhuizen, EM Irusen, and CFN Koegelenberg. 2019. The utility of hand-held mobile spirometer technology in a resource-constrained setting. *SAMJ: South African Medical Journal* 109, 4 (2019), 219–222. https://doi.org/10.7196/SAMJ.2019.v109i4.13845

[11] Julie A Duncan and Paul Aurora. 2014. Monitoring early lung disease in cystic fibrosis: where are we now? *Breathe* 10, 1 (2014), 34–47.

[12] Jerry Alan Fails, Mona Leigh Guha, and Allison Druin. 2013. Methods and techniques for involving children in the design of new technology for children. *Foundations and Trends® in Human–Computer Interaction* 6, 2 (2013), 85–166.

[13] Francis J Gilchrist and Warren Lenney. 2012. The burden of paediatric asthma: economic and familiar. *European Respiratory monograph* 56 (2012).

[14] Global Lung Function Initiative. 2017. Online Calculator v1. http://gligastransfer.org.au/calcs/tlco.html

[15] V Gracchi, M Boel, J Van der Laag, and CK Van der Ent. 2003. Spirometry in young children: should computer-animation programs be used during testing? *European Respiratory Journal* 21, 5 (2003), 872–875.

[16] Brian L Graham, Irene Steenbruggen, Martin R Miller, Igor Z Barjaktarevic, Brendan G Cooper, Graham L Hall, Teal S Hallstrand, David A Kaminsky, Kevin McCarthy, Meredith C McCormack, et al. 2019. Standardization of spirometry 2019 update. An official American thoracic society and European respiratory society technical statement. *American journal of respiratory and critical care medicine* 200, 8 (2019), e70–e88. https://doi.org/10.1164/rccm.201908-1590ST

[17] Murat Gül. 2018. *Konzeption und Design eines Systems zur Diagnose von obstruktiven Lungenerkrankungen bei Kindern.* Dissertation. Technische Universität München, München.

[18] Joep Janssen, Olaf Verschuren, Willem Jan Renger, Jose Ermers, Marjolijn Ketelaar, and Raymond van Ee. 2017. Gamification in physical therapy: More than using games. *Pediatric Physical Therapy* 29, 1 (2017), 95–99.

[19] Matienne Kamp. 2017. *Wearable home-monitoring in asthmatic children.* Master's thesis. University of Twente, Enschede, the Netherlands. Master thesis, https://essay.utwente.nl/73924/.

[20] Maja Kecman, Ifung Lu, Simon Heys, Gianpaolo Fusari, and Matthew James Clear Harrison. 2016. Expiratory flow rate monitoring. European Patent EP3050513A1. Filed January 27, 2015, status withdrawn.

[21] Marije Koopman, Pieter Zanen, Cas LJJ Kruitwagen, Cornelis K van der Ent, and Hubertus GM Arets. 2011. Reference values for paediatric pulmonary function testing: The Utrecht dataset. *Respiratory medicine* 105, 1 (2011), 15–23.

[22] Wanda J Kozlowska and Paul Aurora. 2005. Spirometry in the pre-school age group. *Paediatric respiratory reviews* 6, 4 (2005), 267–272.

[23] Mark L Levy, Philip H Quanjer, Booker Rachel, Brendan G Cooper, Stephen Holmes, and Iain R Small. 2009. Diagnostic Spirometry in Primary Care: Proposed standards for general practice compliant with American Thoracic Society and European Respiratory Society recommendations. *Primary Care Respiratory Journal* 18, 3 (2009), 130–147. https://doi.org/10.4104/pcrj.2009.00054

[24] Joe Marshall, Duncan Rowland, Stefan Rennick Egglestone, Steve Benford, Brendan Walker, and Derek McAuley. 2011. Breath Control of Amusement Rides. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. 73–82. https://doi.org/10.1145/1978942.1978955

[25] Vyaire Medical. 2019. SentrySuite Online Manual v3.0. p22-23, received on request.

[26] Vyaire Medical. 2020. MicroLab Spirometer. Retrieved March 19, 2020 from https://www.vyaire.com/products/microlab-and-microloop-spirometers

[27] Vyaire Medical. 2020. Vyntu SPIRO PC Spirometer. Retrieved March 19, 2020 from https://www.vyaire.com/products/vyntus-spiro-pc-spirometer

[28] MIR Medical International Research. 2015. MIR Medical International Research - English. Retrieved April 20, 2020 from https://vimeo.com/145658007 Points out the tailoring possibility of changing the picture at 0m41s.

[29] MIR Medical International Research. 2020. Spirobank Smart. Retrieved March 19, 2020 from https://www.spirometry.com/eng/Products/spirobank_smart.asp

[30] Virpi Johanna Merikallio, Kirsi Mustalahti, Sami Tapani Remes, Erkka Juhani Valovirta, and Minna Kaila. 2005. Comparison of quality of life between asthmatic and healthy school children. *Pediatric Allergy and Immunology* 16, 4 (2005), 332–340.

[31] Martin R Miller, JATS Hankinson, V Brusasco, F Burgos, R Casaburi, A Coates, R Crapo, P vd Enright, CPM Van der Grinten, P Gustafsson, et al. 2005. Standardisation of spirometry. *European respiratory journal* 26, 2 (2005), 319–338. https://doi.org/10.1183/09031936.05.00034805

[32] Alexander Moeller, Kai-Hakon Carlsen, Peter D Sly, Eugenio Baraldi, Giorgio Piacentini, Ian Pavord, Christiane Lex, and Sejal Saglani. 2015. Monitoring asthma in childhood: lung function, bronchial responsiveness and inflammation. *European Respiratory Review* 24, 136 (2015), 204–215.

[33] Kathleen M Mortimer, Andre Fallot, John R Balmes, and Ira B Tager. 2003. Evaluating the use of a portable spirometer in a study of pediatric asthma. *Chest* 123, 6 (2003), 1899–1907. https://doi.org/10.1378/chest.123.6.1899

[34] Lennart E Nacke and Christoph Sebastian Deterding. 2017. The maturing of gamification research. *Computers in Human Behaviour* (2017), 450–454. https://doi.org/10.1016/j.chb.2016.11.062

[35] Lennart Erik Nacke, Michael Kalyn, Calvin Lough, and Regan Lee Mandryk. 2011. Biofeedback Game Design: Using Direct and Indirect Physiological Control to Enhance Game Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. 103–112. https://doi.org/10.1145/1978942.1978958

[36] Nederlandse Long Alliantie. 2012. Zorgstandaard astma Kinderen & Jongeren [Standards of Care for Asthma in Children and Young People]. http://www.longalliantie.nl/files/9513/7335/4440/Zorgstandaard_Astma_Kinderen_en_Jongeren.pdf.

[37] Nintendo EAD. 2005. *Mario Kart DS.* Game [Nintendo DS]. Nintendo, Kyoto, Japan. Last played 2009.

[38] Nintendo EAD. 2007. *The Legend of Zelda: Phantom Hourglass.* Game [Nintendo DS]. Nintendo, Kyoto, Japan. Last played 2009.

[39] NuvoAir. 2020. Read more about NuvoAir's clinical trial solutions. Retrieved March 19, 2020 from https://www.nuvoair.com/pages/clinical-trial-solutions

[40] W Nystad, SO Samuelsen, P Nafstad, E Edvardsen, T Stensrud, and JJK Jaakkola. 2002. Feasibility of measuring lung function in preschool children. *Thorax* 57, 12 (2002), 1021–1027.

[41] Andreas Oikonomou and David Day. 2012. Using Serious Games to Motivate Children with Cystic Fibrosis to Engage with Mucus Clearance Physiotherapy. In *2012 Sixth International Conference on Complex, Intelligent, and Software Intensive Systems*. 34–39. https://doi.org/10.1109/CISIS.2012.108

[42] Harri Oinas-Kukkonen and Marja Harjumaa. 2009. Persuasive Systems Design: Key Issues, Process Model, and System Features. *Communications of the Association for Information Systems* 24, 28 (2009), 485–500.

[43] Rakesh Patibanda, Florian "Floyd" Mueller, Matevz Leskovsek, and Jonathan Duckworth. 2017. Life Tree: Understanding the Design of Breathing Exercise Games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '17)*. 19–31. https://doi.org/10.1145/3116595.3116621

[44] David E Quinn. 2009. Motivational spirometry system and method. U.S. Patent 7,625,345. Filed March 14, 2005.

[45] Janet C. Read. 2008. Validating the Fun Toolkit: an instrument for measuring children's opinions of technology. *Cognition, Technology & Work* 10, 2 (2008), 119–128.

[46] Janet C. Read and Stuart MacFarlane. 2006. Using the Fun Toolkit and Other Survey Methods to Gather Opinions in Child Computer Interaction. In *Proceedings of the 2006 Conference on Interaction Design and Children*. Association for Computing Machinery, New York, NY, USA, 81–88. https://doi.org/10.1145/1139073.1139096

[47] Paolo Boschetti Sacco. 2009. Incentive Method For The Spirometry Test With Universal Control System Regardless Of Any Chosen Stimulating Image. U.S. Patent 12/298,625. Filed April 26, 2006.

[48] Lamyae Sardi, Ali Idri, and José Luis Fernández-Alemán. 2017. A systematic review of gamification in e-Health. *Journal of biomedical informatics* 71 (2017), 31–48. https://doi.org/10.1016/j.jbi.2017.05.011

[49] Jesse Schell. 2008. *The art of game design: A book of lenses*. Morgan Kaufmann Publishers. 512 pages.

[50] Neuza Silva, Carlos Carona, Carla Crespo, and Maria Cristina Canavarro. 2015. Quality of life in pediatric asthma patients and their parents: a meta-analysis on 20 years of research. *Expert review of pharmacoeconomics & outcomes research* 15, 3 (2015), 499–519.

[51] Tobias Sonne and Mads Møller Jensen. 2016. ChillFish: A Respiration Game for Children with ADHD. In *Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '16)*. 271–278. https://doi.org/10.1145/2839462.2839480

[52] Tobias Sonne, Timothy Merritt, Paul Marshall, Johanne J. Lomholt, Jörg Müller, and Kaj Grønbæk. 2017. Calming Children When Drawing Blood Using Breath-Based Biofeedback. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (Edinburgh, United Kingdom) *(DIS '17)*. Association for Computing Machinery, New York, NY, USA, 725–737. https://doi.org/10.1145/3064663.3064742

[53] Anna Stålberg, Anette Sandberg, Maja Söderbäck, and Thomas Larsson. 2016. The child's perspective as a guiding principle: Young children as co-designers in the design of an interactive application meant to facilitate participation in healthcare situations. *Journal of biomedical informatics* 61 (2016), 149–158.

[54] Paul Tennent, Duncan Rowland, Joe Marshall, Stefan Rennick Egglestone, Alexander Harrison, Zachary Jaime, Brendan Walker, and Steve Benford. 2011. Breathalising Games: Understanding the Potential of Breath Control in Game Interfaces. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology (ACE '11)*. Article 58, 8 pages. https://doi.org/10.1145/2071423.2071496

[55] Rohan Thompson, Ralph J. Delfino, Thomas Tjoa, Eliezer Nussbaum, and Dan Cooper. 2006. Evaluation of daily home spirometry for school children with asthma: New insights. *Pediatric Pulmonology* 41, 9 (2006), 819–828. https://doi.org/10.1002/ppul.20449

[56] Imperial College Healthcare NHS Trust. 2016. Children and young people's - 7 to 12 - Lung function test. Video. Retrieved March 20, 2020 from https://www.youtube.com/watch?v=2knyNHkQ_fg.

[57] Wim M van Aalderen. 2012. Childhood asthma: diagnosis and treatment. *Scientifica* 2012 (2012).

[58] Robby van Delden, Roelof Anne Jelle de Vries, and Dirk K.J. Heylen. 2019. Questioning Our Attitudes and Feelings Towards Persuasive Technology. In *Persuasive Technology: Development of Persuasive and Behavior Change Support Systems (Lecture Notes in Computer Science)*, Harri Oinas-Kukkonen, Khin Than Win, Evangelos Karapanos, Pasi Karppinen, and Eleni Kyza (Eds.). Springer, 3–15. https://doi.org/10.1007/978-3-030-17287-9_1

[59] Robertus Wilhelmus van Delden, Joep Janssen, Silke ter Stal, Wouter Deenik, Winnie Meijer, Dennis Reidsma, and Dirk K.J. Heylen. 2016. Personalization of Gait Rehabilitation Games on a Pressure Sensitive Interactive LED Floor. In *Proceedings of the International Workshop on Personalization in Persuasive Technology co-located with the 11th International Conference on Persuasive Technology (PT 2016) (CEUR workshop proceedings)*, R. Orji, M. Reisinger, M. Busch, A. Dijkstra, A. Stibe, and M. Tscheligi (Eds.). McGill University, 60–73.

[60] Robby van Delden and Koen Vogel. 2020. SpiroPlay. Retrieved July 9, 2020 from https://github.com/hmi-utwente/SpiroPlay

[61] Ruud van der Wel. 2018. Respiratory muscle training. Retrieved March 16, 2020 from https://www.groovtube.nl/en/respiratory-muscle-training-groovtube/

[62] René Van Gent, Liesbeth EM Van Essen, Maroeska M Rovers, Jan LL Kimpen, Cornelis K Van Der Ent, and Gea De Meer. 2007. Quality of life in children with undiagnosed and diagnosed asthma. *European journal of pediatrics* 166, 8 (2007), 843–848.

[63] Daphna Vilozni, Asher Barak, Ori Efrati, Arie Augarten, Chaim Springer, Yacov Yahav, and Lea Bentur. 2005. The Role of Computer Games in Measuring Spirometry in Healthy and "Asthmatic" Preschool Children. *Chest* 128, 3 (2005), 1146 – 1155. https://doi.org/10.1378/chest.128.3.1146

[64] Daphna Vilozni, Michael Barker, Heidemarie Jellouschek, Gerhard Heimann, and Hannah Blau. 2001. An Interactive Computer-Animated System (SpiroGame) Facilitates Spirometry in Preschool Children. *American Journal of Respiratory and Critical Care Medicine* 164, 12 (2001), 2200–2205. https://doi.org/doi:10.1164/ajrccm.164.12.2101002

[65] volksgezondheidenzorg. 2019. Prevalentie based on PIAMA, Preventie en Incidentie van Astma en Mijt Allergie geboortecohort. www.volksgezondheidenzorg.info/onderwerp/astma.

[66] Welch Allyn. 2007. Spiroperfect Spirometry Quick Start Guide. Retrieved April 20, 2020 from https://www.welchallyn.com/content/dam/welchallyn/documents/upload-docs/Training-and-Use/Quick-Reference-Guide/quickguide_20071114_spiroperfect.pdf

[67] Richard Wetzel and Tobias Kreienbühl. 2019. Breathe to Dive: Exploring a Virtual Reality Game for Treatment of Cystic Fibrosis. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 412–416. http://doi.org/10.1109/ISMAR-Adjunct.2019.00044

[68] Windtales. 2019. Lungtraining made fun! Retrieved March 19, 2020 from https://windtales.net/en/

[69] Craig Wright and Elizabeth Conlon. 2009. Critique: Can Children with AD/HD Learn Relaxation and Breathing Techniques through Biofeedback Video Games?. *Australian Journal of Educational & Developmental Psychology* 9 (2009), 47–52.

[70] Zhao Zhao, Ali Arya, Anthony Whitehead, Gerry Chan, and S. Ali Etemad. 2017. Keeping Users Engaged Through Feature Updates: A Long-Term Study of Using Wearable-Based Exergames. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1053–1064. https://doi.org/10.1145/3025453.3025982