Augmented-Reality-Enhanced Product Comparison in Physical Retailing

Jesús Omar Álvarez Márquez
Jürgen Ziegler
jesus.alvarez-marquez@uni-due.de
juergen.ziegler@uni-due.de
Universty of Duisburg-Essen
Duisburg, Germany

ABSTRACT

Augmented reality technology has experienced great improvement in recent years and it has been successfully applied to industry and entertainment settings. However, its application in everyday contexts such as shopping is still very limited. One of the requirements to seamlessly incorporate augmented reality into everyday tasks is to find intuitive, natural methods to make use of it. Due to the inherent capabilities of augmented reality to work as a visual aid to explore and extend the knowledge a user has of the surroundings, this paper proposes the combination of AR technology and product advisors in a novel approach for product comparison. The user's awareness of the differences between multiple physically present objects is enhanced through virtual augmentations, supporting an intuitive way of comparing two or more products while shopping. To assess the validity of the concept, a prototype for an AR-based shopping assistant for comparing vacuum cleaners has been implemented and evaluated in a user study, testing different methods of visual comparison and interaction.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**; **Information visualization**.

KEYWORDS

augmented reality, physical object comparison, comparison visualization, natural interaction

MuC '19, September 8–11, 2019, Hamburg, Germany

ACM ISBN 978-1-4503-7198-8/19/09…$15.00
https://doi.org/10.1145/3340764.3340800

ACM Reference Format:
Jesús Omar Álvarez Márquez and Jürgen Ziegler. 2019. Augmented-Reality-Enhanced Product Comparison in Physical Retailing. In *Mensch und Computer 2019 (MuC '19), September 8–11, 2019, Hamburg, Germany.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3340764.3340800

1 INTRODUCTION AND BACKGROUND

Nowadays, augmented reality (AR) technology is becoming more readily available and holds promise for a multitude of application areas. People's interest on knowing and using the technology has increased too and the number of companies willing to spend resources in adding AR solutions to their products and working processes keeps rising [24, 25]. However, despite the attractiveness of the new technology and its potential to engage consumers [6], the novelty of AR fades away rather quickly [14] and it is difficult to find reasons to use AR technology regularly instead of a more common (and probably convenient) method [15].

On a related note, brick and mortar stores are starting to enhance the shopping experience with the inclusion of computer technologies like smart carts [17], smart shelves [5], RFID sensors [26] or the adoption of the Internet of Things technology [7]. AR has the potential of adding a further quality by bringing the physical and virtual shopping experience together. Especially interesting is its possible application to convey information about physical products or even to work as a personal shop assistant, particularly when the products in question require technical knowledge or the support of an expert opinion to prevent a wrong buying choice. The adoption of AR-based shopping assistants promises to be a beneficial approach for both, retailers and consumers, making the former more competitive and the latter more aware of their buying decisions.

Finding intuitive, natural ways to display and transmit product information requires to take into account current research on customer behaviour. Studies assure that the expected behaviour of a client when in a physical store is to focus on product characteristics, in a decision-making process that requires comparing attributes of different products against each other or against personal preferences [20]. Comparing is one of the most basic cognitive activities and plays an important role in understanding, discovering and evaluating our surroundings [9]. Nonetheless, if the items to be compared are many, retaining their characteristics could impose a big constraint, emphasized by the limitations of short-term memory [3]. In this regard, it has been proven that using spatial superimposition via AR helps to alleviate the mental workload of retaining information by eliminating short-term memory demands [27].

Based on the aforementioned arguments, we present here our approach to physical object comparison through AR technology, where differences between items are highlighted to ease the limitations encountered during the comparison process, aiming to answer the following research questions:

RQ1 What are suitable visualization methods for product comparison?

RQ2 Which interaction style is more convenient for exploring product attributes in a comparative manner?

To validate our approach, a prototype for Microsoft HoloLens has been implemented and evaluated in two different studies. The first one tested two comparison visualizations (total vs relative differences) while the second one focused on interaction methods (tap-based vs head-gaze-based).

2 RELATED WORK

Supporting the comparison of complex data objects has been thoroughly explored from a wide number of perspectives. Gleicher et al. [11] survey on visual comparison research provides a large list of references and establishes a taxonomy of visual designs for comparison, dividing the comparative space in three different categories: juxtaposition (showing different objects separately), superposition (overlaying objects in the same space) and explicit encoding of relationships (computing the relation between objects and visualizing it). Also relevant is the work described in Tominski et al. [28], where a general interaction concept to support comparison tasks in visualization is developed, stressing the great importance of mimicking natural behaviour.

In relation to our interaction concept, head-based techniques have received the interest of the scientific community when applied to 3D environments, especially after being adopted as the standard selection method by popular AR and VR devices (as is the case of the HoloLens or the Oculus Rift). Early work on this area can be seen in Mine [21], where navigation and selection through head movement is included. More recently, Esteves et al. [8] studies the accuracy of head-based input (gaze), proposing a technique for augmented reality based on it and [19] provides an extensive comparison between different multimodal techniques for precision target selection in AR and investigates the combination of eye and head-based tracking.

AR usage in the shopping context has been proposed and studied too. Examples of AR being utilized to enhance the information a client has of the products are the Promopad [29] or the more recent systems presented in Gutiérrez et al. [12] (focussed on providing health-related information of individual items), Ahn et al. [1] (oriented to support product exploration) and Rashid et al. [23] (an approach for browsing physical product shelves).

To the best of our knowledge, visually expressing feature differences and/or similarities of two or more physical objects has not been studied in conjunction with AR yet. Nonetheless, studies dealing with the comparison of physical objects against their digital counterparts could be taken as a reference, as in Georgel et al. [10], where an approach for "discrepancy check" in construction sites is presented.

3 SUPPORTING THE COMPARISON OF PHYSICAL OBJECTS VIA AUGMENTED REALITY

We aim to find suitable interaction and visualization methods for exploring and comparing products in a physical store situation, where augmentations of the available products are provided through a head mounted display (HMD). These augmentations are used to emphasize differences between selected products, making them more noticeable and granting effortless access to information that must otherwise be taken from product flyers or manuals or provided by a human shop assistant. This approach should ease the comparison process, allowing the user to directly visualize the differences of two products regardless of their location within the shop or for which many attributes are available. Thus, it is also within the scope of this research to investigate how such comparison visualization will impact the decision making phase that takes place during the buying process.

Comparison Visualization Methods

Regular comparison means often rely on the use of tables listing product attributes on a side-by-side view, where the information is disconnected from the related physical products. This method may be convenient for online retailers, but it is less than optimal when it comes to brick and mortar stores where customers would need to go back and forth from the real product to its related characteristics. AR, on the other hand, allows for a more direct access to product information, where attributes are shown anchored to the physical object they belong to; then, when the comparison occurs, values from the other compared objects are displayed next to the attributes of the current one, keeping the information attached to the product. In this way, all the information is available even if only one of the products is within the field of view of the customer. Nonetheless, technical products usually have many attributes and filtering techniques may be necessary to prevent overcrowding the display with them, such as the use of attribute categories to only show those chosen by the customer.

Through AR it is possible not only to show superimposed product attributes, but also to link the displayed information to the related parts of the product. For instance, when targeting a vacuum cleaner as a product of interest, it is feasible to visually locate where the dust container is placed. Moreover, numerical data like the dust container's capacity could be enforced by showing a 1:1 3D model of it aligned against the real product, further clarifying the meaning of the number and even directly comparing it, side by side or over the same space, against the capacity of a previously selected item. A main advantage of using AR instead of a traditional approach to product comparison is precisely this: being capable of comparing more than the regular textual or numeric data, but also offering a visualization of what they represent, their exact location, measurements or usage, aspects apt to be compared per se and hardly representable via a different medium.

As mentioned in the related work section, the scientific community has already shown interest regarding the comparison of text and numbers. In our case we mainly make use of a side by side view of these types of values, exploring two different visualization methods (based on the "juxtaposition" and "explicit encodings" categories proposed by Gleicher et al. [11]):

- **Absolute values** Values are presented side by side as they are. Values corresponding to the current item are emphasized, while the values of other selected products are shown next to them, without being modified.
- **Relative values** Values are presented side by side modified depending on how much they differ. Values of the current item are shown unmodified, while the modified values of previously selected items are displayed next to them (e.g. if the current item has a price tag of 50€ and it is being compared against another one that costs 56€, the displayed value will be 50€ and +6€, respectively). When dealing with non numeric values, the performance of the previously selected item regarding this specific attribute is estimated and then an arrow up (better than current item) or down (worse) is used instead (Figure 1).

Figure 1: Absolute vs relative comparison for non-numeric values (for the attribute "filter type").

Interaction Techniques

Performing actions in the digital world trough a HMD may still not feel completely intuitive for a majority of users. People have grown comfortable using traditional user interfaces and conventional interaction mechanisms, thus the importance of making the transition to this new reality as smooth as possible (preventing situations where users could feel lost or incapable to continue without external guidance). However, the use of a HMD allows for a different type of interaction consisting on the activation of virtual elements based on where the user's head is aiming to. The question arises about what method would perform better when exploring the attributes of different products via AR: to make use of those mechanisms users are aware of and utilize with regularity for interacting with digital elements (like the click action) or to employ a technique perhaps more fitting to the nature of a HMD when dealing with real objects (like looking at something). In this regard, our research evaluates two different interaction methods:

- **Explicit activation** This type of interaction makes reference to how users communicate with digital elements by tapping (clicking) on them. In our case, that means that the user will be able to select products, explore and access the different parts of the UI by tapping on holograms or detected real objects.
- **Implicit activation** In a similar fashion to how people show interest for the things around them or inspect the characteristics of a certain object, interaction is carried out through head gaze by pointing at the different UI elements during a dwell time of 0.75 seconds. The value was chosen to be within the limitations regarding application response times, where less than 0.1 seconds feels like the system is reacting instantaneously and more than 1 second may interrupt the user's flow of thought [22].

Prototype

To put this concepts into practice, a prototype AR-based shopping assistant has been implemented for Microsoft's HoloLens platform. A HMD approach was chosen over a different platform because it allows for more interesting interaction possibilities, also leaving the user's hands free to perform a direct inspection of the product. For evaluating purposes the prototype has been conceived to support the comparison of physical vacuum cleaners, although the approach could be easily transposed to different domains. Some of its features are:

Product information visualization There is a distinction between attributes that are linked to a certain part of a product and those that are related to the product in a more general sense. Attributes not related to a specific part are shown floating around it, while the ones referencing parts are "attached" to them. They are organized by the categories "comfort", "performance", "versatility", "maintenance", "filtration" and "accessories" (Figure 2). The categories view displays how well the vacuum cleaner performs on each one of them (using a 1-5 scale), working as a summary of its features. The score given to each category is calculated using a number of custom rules based on how specialized websites assess the quality of the attributes within the category. By selecting one of the categories, the user can access the specific attributes that have an impact on its score (Figure 3).

Attribute explanations In many cases, buying decisions are aggravated when the products have attributes for which expert knowledge is required. To assist costumers in that regard, three different mechanisms have been included, activated when a user selects a single attribute:

- Some attributes act as categories themselves, and selecting them will disclose more specific features. For instance, on the regular view the user can only see (and compare) whether the product includes or not a battery; by selecting the battery attribute other related features will show up, such us battery type, capacity or charging time, allowing for their comparison too (Figure 4).
- A button that displays further information appears, providing a deeper insight about the meaning of the attribute (Figure 5).
- If the attribute is linked to a physical part of the product, such part is highlighted too, showing its shape and location.

Product selection The system knows the morphology of each available product and, once recognized (via the use of markers and the Vuforia SDK), the user can select them by directly tapping on the real object. The system allows the selection of up to three different vacuum cleaners at the same time. Selecting a product will highlight it with a unique colour and include it in the comparison view.

Product comparison When two or three vacuum cleaners are selected at the same time, the comparison view is activated (Figure 3). It comprises the following elements:

- Side by side values: Attributes of the selected products are placed side by side on every chosen vacuum cleaner, distinguishing them by their highlight colours. Values shown will depend on what type of comparison visualization is being used (absolute or relative values).
- Best values: each attribute of the vacuum cleaner which is currently within the field of view is evaluated following the same set of rules used for scoring categories, determining whether it has the best value among the chosen products (highlighted in green) or not (in which case it will appear in red). Besides,

Figure 2: Comparison view of a product in the category section. Current product's ratings are shown in orange, while the ones of a previously selected one appear in blue.

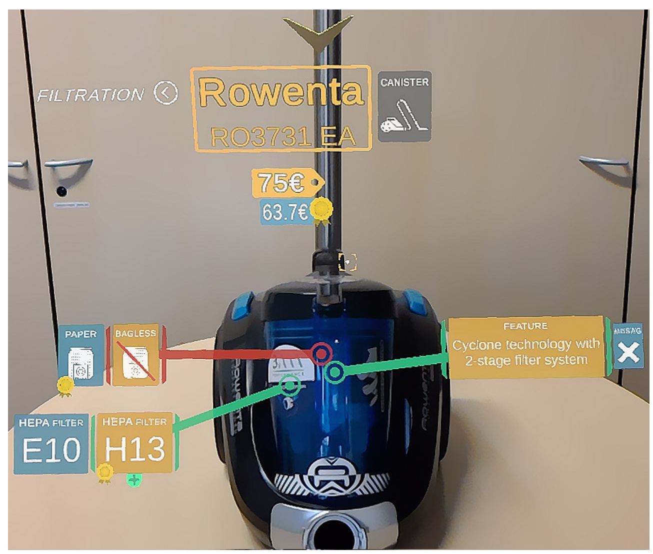

Figure 3: Comparison view of a product within the "filtration" category. Attributes linked to physical parts are attached to them through a line with a circled end.

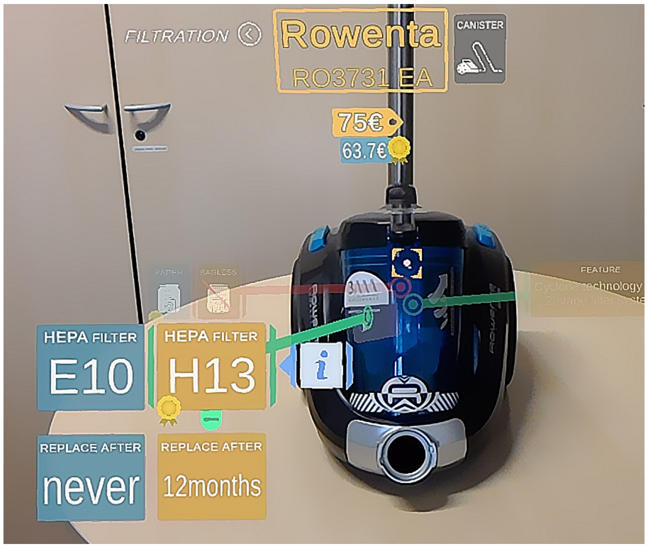

Figure 4: Selecting an attribute opens access to other related attributes (duration of the filter, in this case).

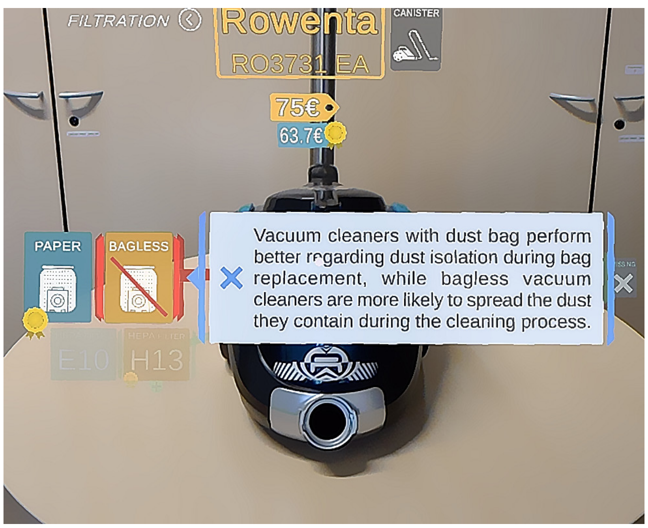

Figure 5: Information relative to the selected attribute.

the best value of a certain attribute among all the selected products appears with a golden badge attached to it.

Interaction The system implements both activation methods previously discussed. Category exploration and attribute selection can both be triggered via implicit or explicit activation, depending on which method is enabled. For vacuum cleaner selection only explicit activation is possible; the use of implicit activation for this purpose was discarded after some testing, which showed that trying to avoid undesired selections required great part of the user's attention.

Table 1: Characteristics of the available products

Feature	VC1	VC2	VC3
Type	handheld and upright	canister	wet-dry
Size	small	medium	big
Weight	light	medium	heavy
Capacity	low	medium	big
Suction power	low	medium	high
Battery	yes	no	no
Action radius	unlimited	large	short
Filter	bad	good	very good
Bag	no	no	yes
Accessories	many	few	many
Price	expensive	average	cheaper
Other	lights	cyclone tech.	blowing

4 EVALUATION

To evaluate our approach, two different studies have been conducted, each of them addressing one of our research questions.

First study: Absolute vs Relative Values

The main objectives of this study were to determine the usability of the approach and to analyse possible significant differences between the two described comparison visualization methods for value representation.

Settings and experimental tasks. Three different vacuum cleaner models were used. They cover different areas of usage but are similar enough so that they can be compared. Their main characteristics are shown in Table 1.

For the study, a floating canvas with instructions was added to the prototype, guiding users through the experiment. On it, three scenarios describing the client's specific needs were presented, asking participants to explore the attributes of the physical vacuum cleaners to find which one would cover the requirements. The given scenarios were:

- Small flat in the city. The tenant has a hairy dog. It also includes the attic, very dark due to the lack of windows.
- Family house with three floors. It has a large backyard with two big trees and lots of dead leaves during autumn. The family's car is usually full of dirt too.
- An elderly person with back problems, living in an old flat where space is scarce. She has a Persian cat and likes gardening in the balcony.

Notice that there was no completely right answer for the available vacuum cleaners in an attempt to increase item exploration. For instance, while the small, battery-powered vacuum cleaner with frontal lights could match most of the requirements for the first scenario, it has poor filtration capabilities which would not be appropriate for a dog owner.

Three versions of the system were implemented, based on different visualization methods for comparing values:

(1) **No comparison enabled (NC)**: a version of the system where comparing items is not available, used as baseline. Participants are still able to visualize and explore product attributes individually.
(2) **Comparison through absolute values (AC)**: comparison is enabled, presenting unmodified values.
(3) **Comparison through relative values (RC)**: comparison is enabled, displaying absolute values for the vacuum cleaner that is within sight, but using relative values for the attributes of the items against this one is being compared.

Method. A total of 50 participants (38 female, average age of 21.16, σ 3.525) took part on the experiment. A between-subjects design was chosen, where only one of the implemented versions of the system was tested by a participant (16 tried NC, 17 AC and 17 RC). Individually, they were taught basic HoloLens usage and interaction possibilities offered by the prototype (for this study only explicit activation was available). After a couple of minutes for letting them get used to it and solve any possible questions they might have, they were told to follow the instructions given by the application and solve the three aforementioned scenarios, which were presented sequentially. After completing all of them, they were given a questionnaire covering aspects related to the ease-of-use of the system. It comprised SUS [4], AttrakDiff [13] and system-specific items measuring the constructs *content quality*, *usefulness* and *future usage intention*. A question directly addressing the preferred kind of visualization for different data types was added too, distinguishing between small numeric, big numeric and non-numeric values.

Results. Regarding the AttrakDiff items, the results can be seen in Table 2 and Figures 6 and 7. All three versions performed very well, falling into the "desired" category, although the one with no comparison received slightly worse results. Similar were the scores reported by the SUS items: 82.81 for the NC version, 83.97 for AC and 85.29 in the case of RC, which qualifies them as "excellent".

Concerning the question about what kind of visualization was preferred depending on the data type, 94% and 96% of the participants chose absolute values for small and big numbers, respectively. For the non-numeric values, 40% selected absolute values, while 60% of them liked the relative comparison more. It has to be noted that the relative comparison visualization in the case of non-numeric values highly differs from the one used for numerals, as previously shown in Figure 1.

Table 2: AttrakDiff's pragmatic (PQ) and hedonic qualities (HQ), along with their respective confidence and the system attractiveness (ATT)*

Version	PQ	Conf.	HQ	Conf.	ATT
No comparison	1.47	0.49	1.44	0.45	1.69
Absolute values	1.75	0.27	1.66	0.41	1.74
Relative values	1.65	0.48	1.74	0.47	1.91

* values provided by http://www.attrakdiff.de

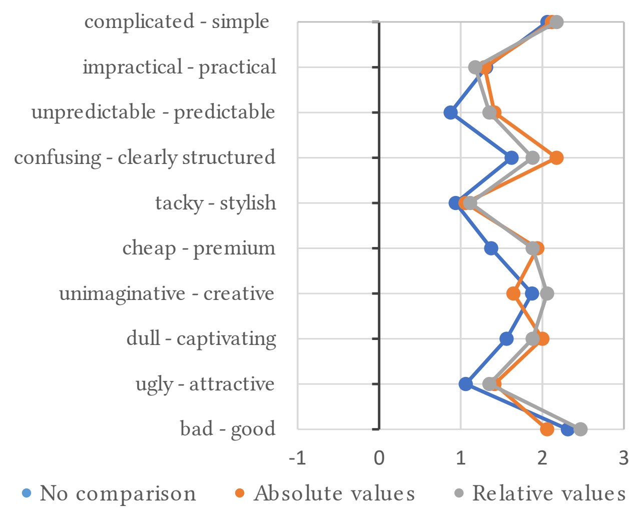

Figure 6: Mean values for AttrakDiff's word pairs in each system version. It uses a -3 to 3 scale, shortened to ease its visualization.

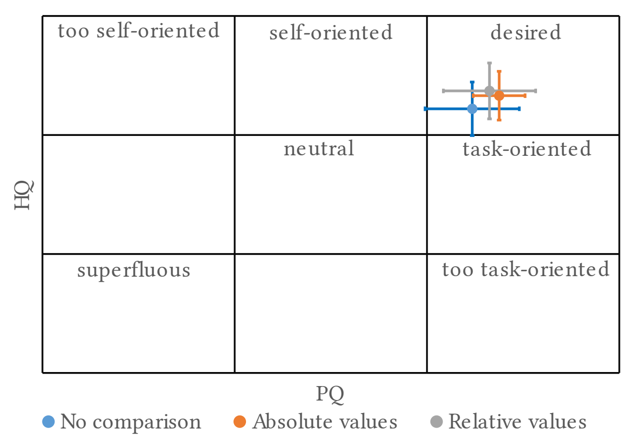

Figure 7: AttrakDiff's pragmatic and hedonic qualities

Table 3: System-specific items

Question	No Comparison		Absolute Values		Relative Values	
	mean	σ	mean	σ	mean	σ
I found the system helpful for deciding for a fitting product.	4.31	1.01	4.35	0.70	4.47	0.62
I found the system helpful for discovering and understanding the product's attributes	4.38	0.61	4.47	0.51	4.65	0.49
The system saved me time in understanding differences between products.	3.38	0.95	3.71	1.10	4.24	1.03
If both options were available, I would prefer to use the AR system instead of a conventional product comparison table	3.19	1.27	3.47	1.23	3.59	1.33

Table 3 shows a list of items covering system-specific aspects not observed by the SUS nor the AttrakDiff questionnaires. A One-way between groups ANOVA analysis was performed over all the items, but there were no significant results.

Discussion. Although there were no significant differences between comparison methods, both of them performed better than the system with no comparison capabilities. It is interesting that despite relative comparison performing slightly better in terms of SUS score and being perceived as more "creative", "good" and "helpful for understanding differences between products", participants would mostly prefer to use absolute comparison instead. It thus seems promising to implement a mixed comparison visualization method that takes into account the type of data to be compared, considering that more than half of the participants chose relative over absolute comparison for non-numeric values, but there was consensus about only using absolute comparison when dealing with numbers. Also, it has to be noted the existence of different, unexplored ways to show relative differences that may have performed better (e.g. the use of percentages) and for which further research is needed. Finally, the novelty of the technology has to be taken into account too, being more than possible that part of the achieved high scores in both SUS and AttrakDiff rely on the fact of using AR.

Second study: Explicit vs Implicit Activation

A second study was performed aiming to explore the possible implications of using implicit vs explicit activation.

Settings and experimental tasks. Mirroring the first study, the same three vacuum cleaners have been used. Equally, participants had to complete the same tasks than before. In relation to our second research question, two new system variations where added based on different input methods:

(1) **Explicit activation (EA)**: activating an attribute and accessing to its information requires a "tap" action.
(2) **Implicit activation (IA)**: using the head to gaze at an attribute activates it, bringing it closer to the user after the dwell time (0.75s) has passed. For accessing its information, the user must aim directly into the pertinent icon. In both cases the pointer's shape will change to display a loading bar, making the user aware of the remaining dwell time before the attribute is selected.

Method. 29 participants took part in the study (17 female, average age of 22.9, σ 3.31). A mixed design was chosen for the experiment, using a within subjects approach for visualization methods and an between-subjects design for activation techniques. To compensate for the impact of the order in which the comparison versions were tested, the conditions where appropriately counterbalanced. By the end, 15 subjects tested EA and 14 IA.

As in the former study, participants where firstly instructed in the usage of the HoloLens. After clarifying any questions, the first task was presented to them and accomplished straight away by using one of the three comparison versions. After its completion, subjects were asked to fill in a short questionnaire. The same procedure was repeated two more times, one per comparison method. Through the whole process, only one activation technique was available.

The questionnaire was composed by the short version of the UEQ proposed in [2] and a set of system-specific items measuring *content quality*, *usefulness* and *future usage intention.*

Results. The results reported by the UEQ questionnaire are presented in Figures 8 and 9. Scores obtained by comparison method confirm the results reported in the previous study, showing values very similar for the three different versions, maintaining the no-comparison one in the worst position. When considering the different methods of activation, IA is perceived as more "creative", "motivating" and "valuable" than EA by a noteworthy difference, while displaying almost equivalent values for the rest of items.

Questionnaire dimensions: A mixed ANOVA repeated measures analysis was performed over the scores obtained for the dimensions measured by the UEQ and the system specific questions (Table 4). The test reported a statistically significant interaction effect between activation and comparison methods for the dimension "hedonic quality" ($F(2, 54) = 5.098, p < .01, \eta_p^2 = .16$). Searching for possible simple main effects, a further multivariate ANOVA test of the between-subjects factor indicated that there is a significant difference in hedonic quality scores between activation methods when using relative comparison ($p < .005$), for which IA obtained better results. Likewise, possible simple main effects of the within-subjects factor were analysed via a one-way repeated measures ANOVA for each activation method separately. For the group using EA, there was a statistically significant effect of the comparison method on hedonic quality scores ($F(2, 28) = 4.485, p < .05, \eta_p^2 = .25$) and performing a Bonferroni test confirmed that there was a significant difference between not using comparison at all and using relative comparison in terms of hedonic quality ($p < .05$), but only when EA is enabled. It has to be noted that, among all the possible combinations, the use of explicit activation with relative values visualization received the lowest score for this dimension. Possible main effects concerning the rest of the dimensions not affected by the interaction between factors were analysed too, but no significant differences were found for any of the independent variables.

Empirical data: Table 5 contains the log data collected during the experiment. A mixed ANOVA repeated measures analysis indicated that there was no statistically significant interaction between activation and comparison methods for any of the measured dependent variables, reason for which the results are omitted here. Equally, no significant main effects were found for comparison methods. Nonetheless, there were significant main effects for activation type, meaning that intervention groups differed significantly, regarding the times an attribute was gazed ($F(1, 27) = 6.49, p < .05, \eta_p^2 = .2$) and tapped ($F(1, 27) = 18.3, p < .001, \eta_p^2 = .41$) as well as how often participants used the help option ($F(1, 27) = 18.42, p < .001, \eta_p^2 = .41$). A closer look to the collected data (Table 5) shows that the times per task were slightly higher for the system without comparison means in both activation techniques, whereas IA shows greater values overall and reaches the highest time per task of all the possible combinations when in conjunction with NC (40% longer than any other time). With EA enabled, participants tended to gaze more into attributes by a great extent. In contrast, during the sessions with IA, subjects selected attributes up to three times more often. Participants' need for visually switching between vacuum cleaners remained fairly consistent in both activation alternatives, although slightly lower values were

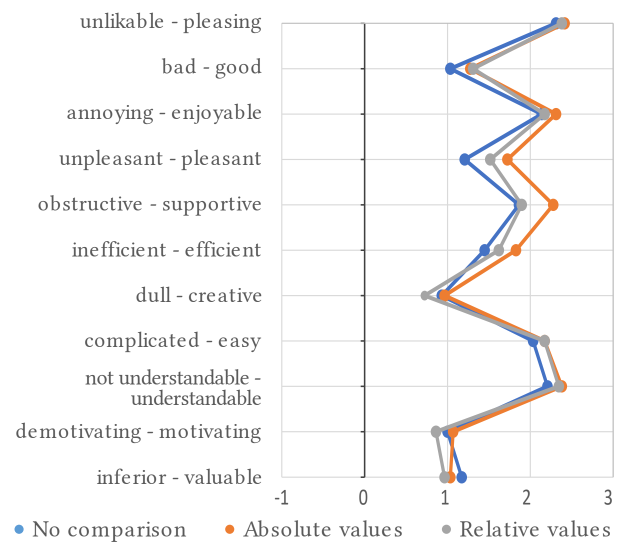

Figure 8: UEQ's items by comparison method. It uses a -3 to 3 scale, shortened to ease its visualization.

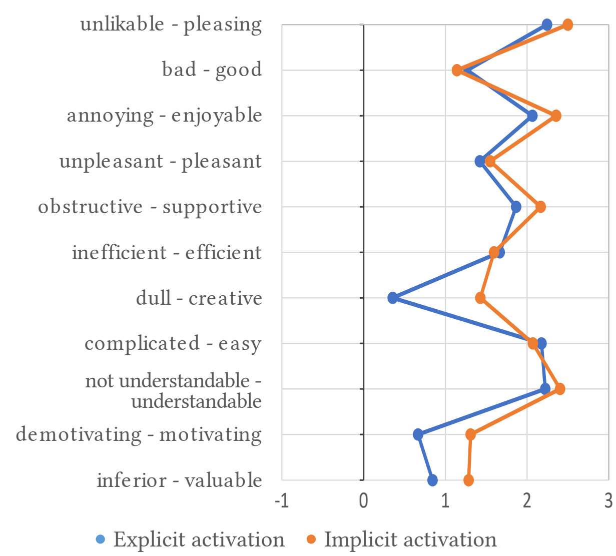

Figure 9: UEQ's items by activation method. It uses a -3 to 3 scale, shortened to ease its visualization.

collected when using RC. On the other hand, the usage of the "information" feature increased greatly with IA and, within this technique, in systems with comparison enabled it was used twice as often than for the no comparison one. Lastly, not all participants made use of the comparison view, preferring to explore each vacuum cleaner and their attributes individually. This is represented by the low percentages of users triggering the comparison, especially when using relative value visualization in conjunction to EA. Comparing three vacuum cleaners at once was a seldom choice in general, showing a modest increment with IA.

Table 4: Mean values for the dimensions measured by the questionnaire. The three first ones are extracted from the UEQ (-3 to 3 scale), while "USE" and "PRE" are system specific (Likert 5 point scale), addressing how useful the prototypes were for exploring/finding products and the participant's preference of usage of the system over a traditional mean, respectively.

	Implicit Activation								Explicit Activation							
	No Comparison		Absolute Comparison		Relative Comparison		Total		No Comparison		Absolute Comparison		Relative Comparison		Total	
	mean	σ	mean	σ	mean	σ	mean	σ	mean	σ	mean	σ	mean	σ	mean	σ
ATT	1.71	0.73	1.98	0.53	1.96	0.61	1.88	0.13	1.63	0.73	1.88	0.62	1.73	0.71	1.75	0.15
PQ	1.96	0.67	2.10	0.57	2.10	0.65	2.06	0.14	1.81	0.79	2.21	0.53	1.91	0.55	1.98	0.12
HQ	1.33	0.70	1.26	0.64	1.42	0.59	1.34	0.15	0.75	0.88	0.80	0.79	0.31	1.13	0.62	0.22
USE†	3.80	0.74	4.04	0.67	4.14	0.93	4.00	0.19	4.04	0.64	4.22	0.55	4.06	0.52	4.11	0.13
PRE†	3.82	0.74	3.96	0.86	3.96	0.81	3.91	0.18	3.91	0.87	3.90	0.76	3.89	0.68	3.90	0.18

† Likert 5 point scale

Table 5: Mean values for the measured empirical factors: required time per task in minutes, number of times a user gazed and tapped (selected) an attribute, how often participants switched their attention from one vacuum cleaner to another, in how many occasions an attribute's information was inspected and the percentage of users who made intentional use of the comparison view of two and three vacuum cleaners (that is, explored items with that view at least during 1 minute).

Version	Time per task	Gaze on attribute	Select attribute	Switch products	Check information	Compare 2 products	Compare 3 products
EA - NC	04:31	152.00	11.47	61.33	3.07	-	-
EA - RC	03:55	132.87	9.13	50.47	0.47	40%	20%
EA - AC	04:28	154.20	11.73	58.87	1.87	80%	20%
IA - NC	07:02	96.00	24.50	62.70	6.64	-	-
IA - RC	04:45	95.85	26.50	60.43	12.21	64%	36%
IA - AC	05:01	90.86	30.86	66.14	11.50	71%	29%

Discussion. The possible benefits of using a comparison technique against not using one are still unclear. In addition to their very similar UEQ results, users appear to visually switch between items with the same frequency no matter whether the comparison is enabled or not, even though the assumption would be to obtain lower values when it is enabled. Although all the pertinent attributes of the item which is out-of-sight are displayed next to the ones belonging to the product the user is currently examining, participants still have the need to visualize the out-of-sight one. Something to consider in this regard is that the products where placed near to each other during the study, which may not be the case in a real world situation. Perhaps the advantages of using the comparison feature would have been more obvious with the added effort of having to walk to a distant product. Testing the system with a set-up closer to a real store may provide further insight about this issue. Despite this, times per task seem to be lower when using a comparison view, suggesting a faster acquisition of information.

Coming back to our first research question, it has to be noted that the comparison view was activated less often when using relative values, probably in association with the results of the previous study where most of the participants chose the absolute values visualization over the relative one. The disinclination to use the relative-values view may be connected to participants having to mentally calculate quantities when applied to numerical data, issue that could be solved by the use of percentages instead of raw numbers. Interestingly, this only seems to affect to the perceived hedonic quality of the system when combined with explicit activation. In general, absolute values visualization seems to be the preferred method for most situations, although the results regarding comparison methods leave open questions and further research is required in this direction, especially considering alternative ways of presenting relative differences and the reasons behind product-switching not being lessened by the usage of the comparison view.

With respect to interaction techniques, the inclusion of implicit activation seems to have a significant effect on the perceived hedonic quality of the system, also reflected in Figure 9, where the three scales in which IA clearly outperforms EA are the ones from which the hedonic quality is calculated. Hence, IA was perceived as more novel or stimulating in general but equally functional than EA (pragmatic quality), probably because IA is a kind of interaction the participants had not used before.

Subjects gazing into attributes significantly less often with IA could be explained by them trying to avoid undesired selections, which would also mean that a more careful navigation was taking place. This works as well as a possible justification for spending more time per task with IA than when using EA in general. It would explain too the very long time per task obtained for the use of implicit activation in a system with no comparison view, specially considering what has been mentioned earlier about the system without comparison means being less effective in terms of information acquisition. Accidentally selecting items is a well studied issue, known as the "Midas Touch" problem [16]. Giving the users awareness about when an attribute is on the course of being selected (like changing the pointer's shape into a timer) and choosing a proper dwell time are helpful methods to mitigate unwanted attribute selections, both of them implemented in the system. Nonetheless, the effects of such issue should still be considered when explaining the higher values for attribute selections during IA. Less probable are the negative effects of the Midas Touch affecting the number of times users accessed the extra information of an attribute: it is a two-step process where first an attribute has to be selected and then the information icon must be "tapped", which is a course of action less susceptible to errors. Given the significant difference between activation methods regarding information checks (also proportionally speaking when compared against attribute selections) it is relatively safe to assume that using IA influenced attribute exploration both in a negative and positive manner, requiring a more careful navigation than EA but at the same time encouraging a more active inspection of the attributes.

The question arises whether the improved hedonic quality and higher attribute inspection when using implicit activation compensate for a harder navigation. As presented in Karapanos et al. [18], the importance of the hedonic quality in terms of the incorporation of a product in daily routines display a sharp decrease after some time has passed, contrary to what happens to the pragmatic quality. This means that the significant difference between hedonic quality scores of both activation methods may not be so after a prolonged usage experience. It is to be expected that participants would have preferred explicit activation in the long run, because it is a less complicated and more direct interaction method and no other significant differences were found. However, it would be interesting to further explore the aspects of implicit activation that encouraged the users to look into the information of the attributes more often and how to take advantage of them when using explicit interaction.

5 CONCLUSIONS AND FURTHER RESEARCH

An approach to product comparison supported by augmented reality has been presented, aiming to enhance the intuitive action of visually comparing two different products when buying in a physical store situation. Augmentations of a product's attributes are shown next to it, also giving further insight about their meaning and highlighting related product parts. Customers can freely select various products, activating a comparison view when two or three are chosen at the same time. The comparison view places the individual attributes of the selected products side by side (emphasizing their differences) and alleviates the effort of remembering them, supporting the decision making process of choosing which one to buy.

A prototype AR-based shopping assistant has been implemented for Microsoft's HoloLens and evaluated in two different user studies. There was a very positive overall outcome in terms of user experience and satisfaction. Results suggest that the inclusion of a comparison feature has a low impact in that regard, although they also indicate a quicker information acquisition when comparison is enabled. When comparing the values of an attribute, their absolute (unmodified) presentation was generally preferred. The implicit selection of attributes through head gaze obtained better results in terms of hedonic quality and attribute examination, but requiring a more careful navigation. Previous studies indicate that in the long run explicit activation may be preferred.

Further research includes looking into the reasons behind the seemingly counter-intuitive results regarding the usage of comparison methods and studying new visualization ways for displaying relative differences. It is also important to identify which aspects of the implicit activation boost the inspection of attributes and find possible ways to apply them to explicit interaction. Additional work is planned for including a recommender system that works on top of the already implemented prototype, giving recommendations not only in the shape of fitting products, but also concerning attributes to be shown (thus, minimizing the need of exploring attributes through the use of categories). It is also in our scope to include a digital catalogue of products, allowing the comparison between physical and non-physical items.

REFERENCES

[1] Junho Ahn, James Williamson, Mike Gartrell, Richard Han, Qin Lv, and Shivakant Mishra. 2015. Supporting Healthy Grocery Shopping via Mobile Augmented Reality. *ACM Trans. Multimedia Comput. Commun. Appl.* 12, 1s (Oct. 2015), 16:1–16:24. https://doi.org/10.1145/2808207

[2] Catherine Alberola, Götz Walter, and Henning Brau. 2018. Creation of a Short Version of the User Experience Questionnaire UEQ. *i-com* 17, 1 (apr 2018), 57–64. https://doi.org/10.1515/icom-2017-0032

[3] G.A. Alvarez and P. Cavanagh. 2004. The Capacity of Visual Short-Term Memory is Set Both by Visual Information Load and by Number of Objects. *Psychological Science* 15, 2 (feb 2004), 106–111.

[4] John Brooke and others. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[5] Steve Brown and David Roth. 2015. The Future of Retail: Shopping and the Smart Shelf. *Intel* (2015). https://www.intel.com/content/www/uk/en/retail/digital-retail-futurecasting-report.html

[6] D. Chatzopoulos, C. Bermejo, Z. Huang, and P. Hui. 2017. Mobile Augmented Reality Survey: From Where We Are to Where We Go. *IEEE Access* PP, 99 (2017), 1–1.

[7] Rong Chen, Li Peng, and Yi Qin. 2010. Supermarket shopping guide system based on Internet of things. In *IET International Conference on Wireless Sensor Network 2010 (IET-WSN 2010)*. IET. https://doi.org/10.1049/cp.2010.1020

[8] Augusto Esteves, David Verweij, Liza Suraiya, Rasel Islam, Youryang Lee, and Ian Oakley. 2017. SmoothMoves: Smooth Pursuits Head Movements for Augmented Reality. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, New York, NY, USA, 167–178. https://doi.org/10.1145/3126594.3126616 event-place: Québec City, QC, Canada.

[9] Dedre Gentner and José Medina. 1997. Comparison and the Development of Cognition and Language. *Cognitive Studies* 4, 1 (March 1997), 1_112–1_149.

[10] Pierre Georgel, Pierre Schroeder, Selim Benhimane, Stefan Hinterstoisser, Mirko Appel, and Nassir Navab. 2007. An Industrial Augmented Reality Solution For Discrepancy Check. In *Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR '07)*. IEEE Computer Society, Washington, DC, USA, 1–4. https://doi.org/10.1109/ISMAR.2007.4538834

[11] Michael Gleicher, Danielle Albers, Rick Walker, Ilir Jusufi, Charles D. Hansen, and Jonathan C. Roberts. 2011. Visual comparison for information visualization. *Information Visualization* 10, 4 (sep 2011), 289–309.

[12] Francisco Gutiérrez, Katrien Verbert, and Nyi Nyi Htun. 2018. PHARA: An Augmented Reality Grocery Store Assistant. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '18)*. ACM, New York, NY, USA, 339–345. https://doi.org/10.1145/3236112.3236161

[13] Marc Hassenzahl, Michael Burmester, and Franz Koller. 2003. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In *Mensch & Computer 2003: Interaktion in Bewegung*, Gerd Szwillus and Jürgen Ziegler (Eds.). Vieweg+Teubner Verlag, Wiesbaden, 187–196. https://doi.org/10.1007/978-3-322-80058-9_19

[14] Toby Hopp and Harsha Gangadharbatla. 2016. Novelty Effects in Augmented Reality Advertising Environments: The Influence of Exposure Time and Self-Efficacy. *Journal of Current Issues & Research in Advertising* 37, 2 (July 2016), 113–130.

[15] Tseng-Lung Huang and Shuling Liao. 2015. A model of acceptance of augmented-reality interactive technology: the moderating role of cognitive innovativeness. *Electronic Commerce Research* 15, 2 (June 2015), 269–295.

[16] R Jakob. 1998. The use of eye movements in human-computer interaction techniques: what you look at is what you get. *Readings in intelligent user interfaces* (1998), 65–83.

[17] Gerrit Kahl, Lübomira Spassova, Johannes Schöning, Sven Gehring, and Antonio Krüger. 2011. IRL SmartCart - a User-adaptive Context-aware Interface for Shopping Assistance. In *Proceedings of the 16th International Conference on Intelligent User Interfaces (IUI '11)*. ACM, New York, NY, USA, 359–362. https://doi.org/10.1145/1943403.1943465 event-place: Palo Alto, CA, USA.

[18] Evangelos Karapanos, John Zimmerman, Jodi Forlizzi, and Jean-Bernard Martens. 2009. User experience over time. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. ACM Press. https://doi.org/10.1145/1518701.1518814

[19] Mikko Kytö, Barrett Ens, Thammathip Piumsomboon, Gun A. Lee, and Mark Billinghurst. 2018. Pinpointing: Precise Head- and Eye-Based Target Selection for Augmented Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, 81:1–81:14. https://doi.org/10.1145/3173574.3173655 event-place: Montreal QC, Canada.

[20] Kelvin J. Lancaster. 1966. A New Approach to Consumer Theory. *Journal of Political Economy* 74, 2 (April 1966), 132–157.

[21] Mark R. Mine. 1995. *Virtual Environment Interaction Techniques*. Technical Report. University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

[22] Jakob Nielsen. 1994. *Usability engineering*. Elsevier.

[23] Zulqarnain Rashid, Rafael Pous, Joan Melià-Seguí, and Marc Morenza-Cinos. 2014. Mobile Augmented Reality for Browsing Physical Spaces. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp '14 Adjunct)*. ACM, New York, NY, USA, 155–158. https://doi.org/10.1145/2638728.2638796 event-place: Seattle, Washington.

[24] Alexandra Rese, Daniel Baier, Andreas Geyer-Schulz, and Stefanie Schreiber. 2017. How augmented reality apps are accepted by consumers: A comparative analysis using scales and opinions. *Technological Forecasting and Social Change* 124 (Nov. 2017), 306–319. https://doi.org/10.1016/j.techfore.2016.10.010

[25] Lucas Roitman, Jeff Shrager, and Terry Winograd. 2017. A Comparative Analysis of Augmented Reality Technologies and their Marketability in the Consumer Electronics Segment. *Journal of Biosensors & Bioelectronics* 08, 01 (2017). https://doi.org/10.4172/2155-6210.1000236

[26] Michael Schneider. 2003. A Smart Shopping Assistant utilising Adaptive Plan Recognition. *ABIS 2003 – 11. GI-Workshop "Adaptivität und Benutzermodellierung in interaktiven Softwaresystemen"* (2003). http://dl.gi.de/handle/20.500.12116/5123

[27] Arthur Tang, Charles Owen, Frank Biocca, and Weimin Mou. 2003. Comparative effectiveness of augmented reality in object assembly. In *Proceedings of the conference on Human factors in computing systems - CHI '03*. ACM Press.

[28] C. Tominski, C. Forsell, and J. Johansson. 2012. Interaction Support for Visual Comparison Inspired by Natural Behavior. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (dec 2012), 2719–2728.

[29] Wei Zhu and Charles B. Owen. 2008. Design of the PromoPad: An Automated Augmented-Reality Shopping Assistant. *JOEUC* 20 (2008), 41–56.